The development and use of *in vitro* and in silico Beta-site Amyloid Precursor Protein cleaving enzyme-1 protein-ligand binding assays

by

Mark John Chambers

Submitted in fulfilment of the academic requirements of

Doctor of Philosophy

in Biochemistry

School of Life Science

College of Agriculture, Engineering and Science

University of KwaZulu-Natal

Pietermaritzburg

South Africa

2020

PREFACE

The research contained in this dissertation was completed by the candidate while based in the Discipline of Biochemistry, School of Life Science of the College of Agriculture, Engineering and Science, University of KwaZulu-Natal, Pietermaritzburg, South Africa. The research was financially supported by the National Research Foundation of South Africa and the Transdisciplinary Flagship program: APACHE (Afrocentric Precision Approach to Control Health Epidemics).

The contents of this work have not been submitted in any form to another university and, except where the work of others is acknowledged in the text, the results reported are due to investigations by the candidate.

Signed: Dr. Raymond Hewer

Date: November 2020

DECLARATION 1: PLAGIARISM

I, Mark John Chambers, declare that:

(i)　　the research reported in this dissertation, except where otherwise indicated or acknowledged, is my original work;

(ii)　　this dissertation has not been submitted in full or in part for any degree or examination to any other university;

(iii)　　this dissertation does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons;

(iv)　　this dissertation does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

　　　　a)　　their words have been re-written but the general information attributed to them has been referenced;

　　　　b)　　where their exact words have been used, their writing has been placed inside quotation marks, and referenced;

(v)　　where I have used material for which publications followed, I have indicated in detail my role in the work;

(vi)　　this dissertation is primarily a collection of material, prepared by myself, published as journal articles or presented as a poster and oral presentations at conferences. In some cases, additional material has been included;

(vii)　　this dissertation does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the dissertation and in the References sections.

_____

Signed: Mark John Chambers

Date: November 2020

DECLARATION 2: PUBLICATIONS

My role in each paper and presentation is indicated. The * indicates corresponding author.

1. Chambers, M. J. Hewer, R.* 2018. The expression and purification of β-site APP cleaving enzyme-1 (BACE1) in *Escherichia coli.* Poster presentation to the South African Society for Biochemistry and Molecular Biology, 8th to 10[th-] July, 2018, Potchefstroom, South Africa. Presented and compiled by Chambers, M. J.

2. Chambers, M. J. Delport, A. Hewer, R.* 2020. The use of the cellular thermal shift assay for detection of intracellular beta-site amyloid precursor protein cleaving enzyme-1 ligand binding. Article submitted to Molecular Biology Reports. Paper written and experiments carried out by Chambers M. J.

_____ ████████████_____

Signed: Mark John Chambers

Date: November 2020

ABSTRACT

The β-site APP cleaving enzyme-1 (BACE1) is a transmembrane aspartic protease produced in humans, being most active in the human brain. BACE1 is heavily implicated in the Alzheimer's disease pathology where it mediates the rate-limiting step in the production of amyloid-β peptide. Expression and purification of BACE1 is associated with low yields due to aggregation as a result of its inherent insolubility. In this study, expression of a recombinant form of BACE1 in *Escherichia coli* is described. This expression was successfully optimized to generate yields of 12 mg/L of culture, with the optimized protocol involving induction using 1 mM IPTG in the presence of 1% ethanol overnight at 16 ˚C. The expressed protein was demonstrated to be insoluble after induction and was resolubilized and refolded using a 24-hour incubation in a 10% N-lauryl sarcosine sodium salt solution. Refolding was confirmed through the use of a gelatine gel zymogram which indicated that the protein was catalytically active. The use of gelatine gel zymograms for assessment of BACE1 activity have not previously been described in literature. The purified recombinant BACE1 was confirmed to be immunogenic in chickens with the purified antibodies being capable of successfully detecting BACE1 produced in mammalian cells. A differential scanning fluorimetry (DSF) assay was then optimized using the recombinant BACE1, in an effort to screen for BACE1 binding compounds. The DSF assay was carried out in phosphate buffered saline (PBS), pH 7.4 at a concentration of 1 μM with the unfolding being confirmed to occur without the formation of a defined unfolding intermediate, exhibiting a melting temperature (Tm) of 82 ± 0.21°C. The standard Gibbs free energy, entropy and enthalpy of unfolding of BACE1 in PBS were calculated to be 112 ± 0.1 kJ/mol, 1.82 ± 0.002 kJ/mol K and 642 ± 0.7 kJ/mol respectively. In the presence of the BACE1 specific inhibitor, Verubecestat, the Tm rose to 84.15 ± 0.28°C. An in-house compound library derived from the ChemBridge DiverSet compound library was then screened against BACE1 using this DSF assay. Three compounds were found to induce a thermal shift, indicating binding. These compounds were N-ethyl-N',N'-dimethyl-N-[2-(trifluoromethyl)benzyl]-1,2-ethanediamine (**C19**) displaying a ΔTm of 2.4 0.65 ˚C, 3-cyclopentyl-N-(4-pyridinylmethyl)propanamide (**C34**) with a ΔTm of 3.045 ± 0.66 ˚C and 4-chloro-1-(2-ethoxybenzoyl)-1H-pyrazole (**C39**) with a ΔTm of 2.625 ± 0.64 ˚C. The in-house compound library was also screened, in silico, against a 3D BACE1 model which illustrated that both **C19** and **C34** docked to the catalytic residues of BACE1. It was established that **C19** likely bound a non-catalytic binding site as well. No appreciable docking was detected for the **C39** compound in silico however, indicating the possibility of an undetected binding site. Using a BACE2 in silico model, the specificity of

C34 for BACE1 over BACE2 was also predicted. Intracellular BACE1 binding capability of each test compound was assessed through the first reported use of a cellular thermal shift assay for screening for BACE1 binding compounds. The cellular thermal shift assay was carried out using the known BACE1 inhibitor Verubecestat, where an increase in T_{agg} to $53.27 \pm 0.89°C$ from $49.53 \pm 0.69°C$ was observed. The three test compounds were then screened with **C34** yielding a significant increase in T_{agg} (p-value < 0.05) from 53.09 ± 1.30 confirming intracellular binding. Compound **C34** was then used in an iterative screening of the Enamine Real compound database for compound analogues using the in silico BACE1 model, ultimately resulting in the identification of compound **C5064**. Compound **C5064** was detected to bind to the BACE1 catalytic site yielding a Glide score of -10.121. The cellular thermal shift assay was again used, confirming **C5064** possessed the ability to bind intracellular BACE, resulting in the second novel intracellular BACE1 binding molecule identified in this study. Finally, a zebrafish embryo toxicity test was carried out with **C19**, **C34** and **C39** generating LC_{50} values of 26.125 ± 2.45, 127 ± 9.9 and 4.52 ± 0.61 µM respectively. Only **C34** was detected to be teratogenic yielding a teratogenicity index of 1.33. No LC_{50} concentration or teratogenicity index was calculated for compound **C5064** however, due to the compound exhibiting no lethality within its solubility range. This study succeeded in thoroughly assessing the thermostability of BACE1, both recombinantly expressed in *Escherichia coli* as well as in mammalian human embryonic kidney-293 cells. Four novel BACE1 binding compounds were detected with two confirmed to bind BACE1 intracellularly, with the first cellular thermal shift assay carried out on BACE1 being reported. Finally, all test compounds were successfully assessed for toxicity and teratogenicity *in vivo* using a zebrafish embryo toxicity test.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

$\Delta_u G$	Gibbs free energy of unfolding
$\Delta_u H$	Enthalpy of unfolding
$\Delta_u S$	Entropy of unfolding
2xYT	2 x yeast tryptone
3D	three dimensional
AIDS	acquired immunodeficiency syndrome
$A\beta$	amyloid-β peptide
AchE	acetylcholinesterase
AD	Alzheimer's disease
ADAM9	a disintegrin and metalloproteinase 9
ADAM10	a disintegrin and metalloproteinase 10
ADAM17	a disintegrin and metalloproteinase 17
AICD	App intracellular domain
AMT	acetate, MES, Tris-HCl
APP	amyloid precursor protein
BACE1	β-site APP-cleaving enzyme-1
BACE2	β-site APP-cleaving enzyme-2
Bp	base pairs
BSA	bovine serum albumin
CC_{50}	50% of the maximum cytotoxicity concentration
CETSA	cellular thermal shift assay
CHO	Chinese hamster ovary
DMEM	Dulbecco's Modified Eagle's Medium
DSF	differential scanning fluorimetry
E. coli	Escherichia coli
ECL	enhanced chemiluminescence
EC_{50}	concentration required to reduce the morphological score to 50%
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbent assay
FBS	foetal bovine serum
FDA	Food and Drug administration

FRET	**fluorescence resonance energy transfer**
Glide	**Grid-based Ligand Docking with Energetics**
GST	**glutathione S-transferase**
HEK	**human embryonic kidney**
HEPES	**4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid**
Hpf	**Hours post-fertilization**
HRM	**high resolution melting**
HRPO	**horse radish peroxidase**
IC_{50}	**50% of the maximum inhibitory concentration**
IPTG	**isopropyl β-D-1-thiogalactopyranoside**
Ki	**inhibitor constant**
LB	**lysogeny broth**
LC_{50}	**50% of the lethal concentration**
MES	**2-(N-Morpholino) ethanesulfonic acid**
MTS	**3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium salt**
MWCO	**molecular weight cut-off**
MWM	**molecular weight marker**
NFT	**neurofibrillary tangles**
NIAID	**National Institute of Allergy and Infectious Diseases**
NIH	**National Institute of Health**
NMDAR	**N-Methyl-D-aspartate receptor**
NMR	**nuclear magnetic resonance**
OD	**optical density**
PAGE	**polyacrylamide gel electrophoresis**
PBS	**phosphate-buffered saline**
PDB	**Protein DataBank**
PEG	**polyethylene glycol**
PenStrep	**penicillin-streptomycin**
PES	**phenazine ethosulphate**
PS	**presenilin**
PVDF	**polyvinylidene fluoride**
QSAR	**quantitative structure-activity relationship**
R^2	**coefficient of determination**

RIPA	**radioimmunoprecipitation assay**
RMSD	**root-mean-square-deviation**
ROS	**reactive oxygen species**
RPM	**revolutions per min**
RSA	**Republic of South Africa**
sAPPα	**secreted α-amyloid precursor protein**
sAPPβ	**secreted β-amyloid precursor protein**
Sarkosyl	**N-lauryl sarcosine sodium salt**
SDS	**sodium dodecyl sulphate**
SOC	**super optimal broth with catabolite repression**
SP	**standard precision**
TAE	**Tris-HCl, acetic acid, EDTA**
T$_{agg}$	**apparent aggregation temperature**
TBS	**Tris-buffered saline**
TI	**teratogenicity index**
Tm	**melting temperature**
TMB	**3,3',5,5'-Tetramethylbenzidine**
Tris	**tris (hydroxymethyl) aminomethane**
USA	**United States of America**
XP	**extra precision**
ZFET	**zebrafish embryo toxicity test**

CHAPTER 1: LITERATURE REVIEW

1.1 Introduction

The first description of Alzheimer's disease (AD) was made by the German psychiatrist Alois Alzheimer in a lecture delivered in 1906. Here he described a disease he had noticed in a 55-year-old patient, which resulted in cognitive decline followed by death. It was discovered through post-mortem examination that senile amyloid plaques, alongside neurofibrillary tangles (NFT), were present in the brain of the patient. Up until then senile plaques had only been documented in the elderly, while the NFTs were a novel discovery at the time. Alzheimer would publish an article describing this briefly in 1907 and later again in 1911 in which he further elucidated on his findings (Alzheimer, 1907, Alzheimer, 1911).

It was only in 1984 however, that the first successful purification of the amyloid protein which comprises senile plaques was carried out by Glenner and Wong (Glenner and Wong, 1984). The peptide was designated amyloid-β (Aβ) due to its β-pleated sheet structure. Description of the composition of NFTs followed and in 1986 it was reported, through the use of antibodies raised against NFTs, that the microtubule-associated tau protein was a major component thereof, with ubiquitin being another (Nukina and Ihara, 1986; Mori *et al*., 1987). After the first successful production of cDNA encoding the amyloid precursor protein (APP) gene in 1987, several gene mutations were associated with AD susceptibility (Kang *et al*., 1987). These mutations were found in genes for presenilin, APP, as well as apolipoprotein E4 (Corder *et al*., 1993, Citron *et al*., 1992, Sherrington *et al*., 1996). In 1993 the first drug against AD was approved by the Food and Drug Administration of the United States of America (FDA). Since then only four other treatments have been approved for commercial use against AD. AD is currently the most common form of dementia (Sardana *et al*., 2003). AD susceptibility has been correlated most strongly with age, however it has also been linked to various lifestyle choices such as smoking and poor diet (Ott *et al*., 1998; Engelhart *et al*., 2012).

1.2 Pathophysiology of Alzheimer's disease

1.2.1   Amyloid hypothesis pathology

The broad pathophysiology of AD involves four major pathways (Cole and Vassar, 2007). Namely the amyloid, tau, cholinergic and inflammatory pathways, each of which is implicated in the characteristic cognitive decline in AD (Kar *et al*., 2002; Gong and Iqbal, 2008). According to the amyloid hypothesis, Aβ plays a role in initiating all four pathways. The hypothesis proposes that when a neuron generates excess Aβ, this Aβ will bind to the α-7 nicotinic receptors on neighbouring astrocytes (Pirttimaki *et al*., 2013). This binding action induces the secretion of glutamate*via*exocytosis, which in turn binds to N-Methyl-D-aspartate receptors (NMDARs) found on the original neuron, which results in an uptake of Ca^{2+} ions into the cell (Shankar *et al*., 2007). The imbalance in Ca^{2+} ions causes mitochondria to release reactive oxygen species (ROS) which bring about oxidative stress (Reynolds and Hastings, 1995). Both the influx of calcium as well as the oxidative stress are thought to dissociate tau from the neurons microtubules, allowing for its hyperphosphorylation by various kinases (Su *et al*., 2010) (Mattson *et al*., 1991). The hyperphosphorylated tau then aggregates inside the neuron forming NFTs (Nukina and Ihara, 1986). These NFTs cause cellular damage through actin destabilization, mitochondrial damage as well as through reducing the ability of a neuron to transfer neurotransmitters through synapses (Kraemer *et al*., 2003). The tau oligomers that form these NFTs also spread to other neurons through synapses, after which they interact with unaltered tau converting it into its amyloid-like form (Guo and Lee, 2011). Excess Aβ also forms extracellular plaques which induce inflammation through the action of microglia (Sasaki *et al*., 2002). This pro-inflammatory response of microglia to Aβ plaques results in the secretion of ROS, as well as several inflammatory cytokines such as interleukin-1β, tumour necrosis factor α and interferon γ (Wilkinson and Landreth, 2006; Wang *et al*., 2015). The resulting combination of NFTs, Ca^{2+} imbalance, inflammation and amyloid load results in a proliferative cycle, causing neurodegeneration in the process. This damage to neurons results in cognitive decline, most notably when cholinergic neurons are damaged as this disrupts the brains ability to produce the neurotransmitter acetylcholine, which plays roles in memory and learning (Rasool *et al*., 1986). This deficit of acetylcholine is further compounded by the ability of Aβ to induce acetylcholine degradation*via*cholinesterase enzymes (Kumar *et al*., 2018).

1.2.2 Concerns regarding the amyloid hypothesis

The amyloid hypothesis is complicated by the identification that amyloid plaques are present in 20-40% of the elderly population that do not suffer from the disease and often in high enough amounts so as to warrant an AD diagnosis (Drachman, 2014). Originally it was assumed that at a certain level of amyloid burden the cognitive decline was induced, however the fact that such large portions of the elderly population exhibit high amyloid load without Alzheimer's symptoms made this appear unlikely. Further, it is difficult to determine whether in these cases the symptoms simply have yet to present themselves, as this would require constant mental evaluation (Kepp, 2016). The amyloid hypothesis was altered to accommodate this based on the presence of various subtypes of the Aβ peptide, most important of which are the Aβ40 and Aβ42. The modified amyloid hypothesis now assumes that, rather than an overall Aβ load causing cognitive decline, it is, in fact, the ratio between these two forms of the Aβ peptide that instigates it, a theory supported by the knowledge that Aβ42 is substantially more toxic than Aβ4. This is further supported by the fact that several presenilin mutations affect this ratio and have indeed been correlated to the onset of AD (Tiwari and Kepp, 2015). There still remains however, speculation around how exactly this ratio results in the cognitive decline. Another area where further research is required involves the location at which the amyloid plaques occur. AD is associated with degradation of specific regions of the brain, most notably the hippocampus and amygdala, whereas the location of amyloid plaques do not correspond as they are often distributed across the entire brain before AD symptoms present themselves (Kepp, 2016). This, along with the knowledge that level of amyloid load has not been correlated to cognitive decline causes concern for this theory (Drachman, 2014).

1.2.3 Generation of amyloid-β

The Aβ peptide is generated from the sequential cleavage of APP, it can however be cleaved through one of two pathways, the amyloidogenic and non-amyloidogenic pathways, summarized in Figure 1.1. In the non-amyloidogenic pathway, APP is cleaved by a group of proteases, known as α-secretases, which includes the enzymes, a disintegrin and metalloproteinase domain-containing proteins 9, 10 and 17 (ADAM9, ADAM10 and ADAM17 respectively) with ADAM10 being the primary α-secretase (Kuhn *et al*., 2010). This occurs on the cell surface. In the amyloidogenic pathway, as shown in Figure 1.1, APP is cleaved sequentially by two enzymes. First the β-secretase, an aspartic protease named β-site APP cleaving enzyme-1 (BACE1) cleaves the APP at its β-cleavage site (Cai *et al*., 2001). This

typically occurs in the endosomes and Golgi apparatus after APP is internalized into the cell and results in two products, namely secreted β-amyloid precursor protein (sAPPβ) and C99 (Yan, 2017). The transmembrane C99 is then cleaved by the protein complex ɣ-secretase comprised of presenilin 1 or 2, nicastrin, presenilin enhancer-2 and anterior pharynx defective-1 (Zhang *et al*., 2001; Esler *et al*., 2002; Kimberly *et al*., 2003). This results in the formation of the APP intracellular domain (AICD) as well as the Aβ peptide.
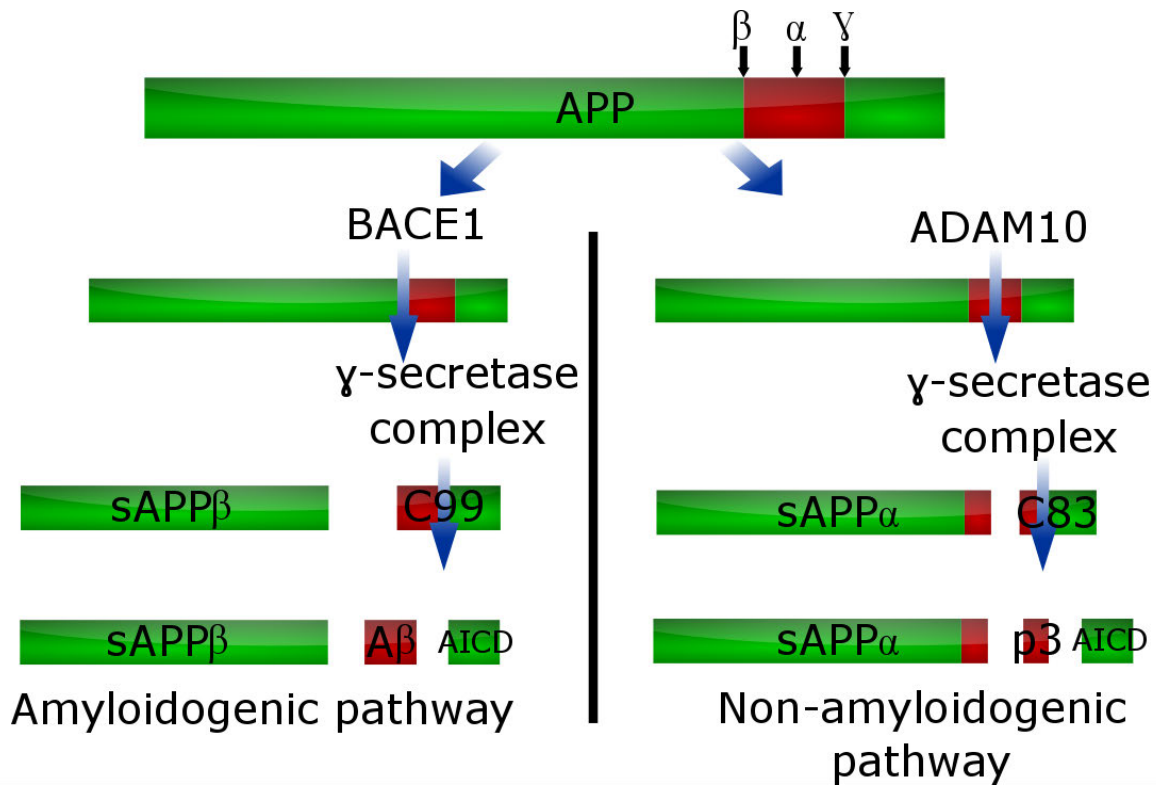
Figure 1.1: The diverging pathways of amyloid precursor protein (APP) processing by the secretase-class enzymes. APP contains three cleavage sites, each of which is recognized by either β, α or ɣ-secretase. In the amyloidogenic pathway, the β-secretase, β-site APP cleaving enzyme-1 (BACE1), cleaves APP first, at the β-cleavage site producing secreted β-amyloid precursor protein (sAPPβ) and C99. After this, the ɣ-secretase complex cleaves the C99 peptide into either Amyloid β (Aβ) and the APP intracellular domain (AICD). In the non-amyloidogenic pathway, α-secretase cleaves APP, first generating secreted α-amyloid precursor protein (sAPPα) and C83, after which, ɣ-secretase cleaves the C83 peptide into the peptide p3 and AICD. Adapted from Prete *et al*., 2014.

1.2.4 *Intervention strategies*

Some attempts have been made, in mouse models to modulate the activity of the α-secretase enzymes in order to reduce the amount of Aβ peptide being produced (Postina *et al*., 2004). Through this, overexpression of ADAM10 resulted in the majority of APP being processed in

the non-amyloidogenic pathway which resulted in a decrease in cognitive decline. Although no clinical trials have as of yet been carried out, it was described in 2018 that the compound cosmosiin increased ADAM10 expression in neuroblastoma cells (Min *et al*., 2018). With regards to direct inhibition of the amyloidogenic pathway, ɣ-secretase inhibition was a possible target. Side-effects of its inhibition are of concern with regards to human treatment, as the complex is known to process not only APP but also mediate the cleavage of proteins such as Notch, E-cadherin and CD44 (Marambaud *et al*., 2002; Kimberly *et al*., 2003; Murakami *et al*., 2003). This problem was highlighted in the trials of Semagacestat, the ɣ-secretase inhibitor from Eli Lilly (Hopkins, 2011). Inhibition of the ɣ-secretase enzyme resulted in increased cognitive decline as well as an increased incidence of skin cancer in patients (De Strooper, 2014).

Another avenue for direct inhibition of the amyloidogenic pathway is BACE1 inhibition. Cleavage of APP by BACE1 has not only been demonstrated to be the rate-limiting step of Aβ generation (Stockley and O'Neill, 2007) but, in a study on BACE1-knockout mice (Luo *et al*., 2003) a lack of BACE1 did indeed reduce the amount of Aβ present in the adult mice. Concern for the use of BACE1 inhibition as a method of AD treatment is directed towards the lack of knowledge surrounding its functions outside of APP cleavage as BACE1 knockout mice, although not presenting with AD, do present some abnormal phenotypes such as seizures resulting from an over excitation of neurons as well as hypomyelination (Hu *et al*., 2006). Further, with the recent failure of the BACE1 inhibitors, Verubecestat and Lanabecestat (Hawkes, 2017; Burki, 2018), it has become apparent that these BACE1 inhibitors resulted in an increase in the rate of cognitive decline in patients, this was not seen in the phase 2 trials of Elenbecestat, another BACE1 inhibitor produced by Biogen, Eisai Co., Ltd., however. BACE1 has been implicated in several brain functions such as neuronal plasticity, synapse formation and myelination which could explain the adverse effects of BACE1 inhibition. Several studies have tried to determine possible BACE1 substrates in an effort to determine what adverse effects could result from its inhibition (Hemming *et al*., 2009; Kuhn *et al*., 2012) . These studies suggest that BACE1 might play a role in neurogenesis (Modarresi *et al*., 2011). Further efforts to generate AD drugs have followed the exploration of several other modes of treatment, including the use of antibodies against either amyloid-β or tau protein (Doody *et al*., 2014; Bohrman *et al*., 2012; Cummings *et al*., 2014; Sevigny *et al*., 2016). The goal of these antibodies is to illicit Aβ phagocytosis without causing an inflammatory response.

The first protein targeted as a potentially important component of AD pathology was the enzyme acetylcholinesterase (AchE), with three of the four currently marketed AD drugs being AchE inhibitors (Tabet, 2006). Inhibition of AchE reduces the rate at which acetylcholine is degraded, thereby maintaining the levels of the neurotransmitter in the brain, slowing cognitive decline. The fourth drug, and the only FDA approved treatment method that does not target acetylcholinesterase is an NMDAR inhibitor in the glutamatergic system.

1.3 Targeting of Alzheimer's disease pathways over proteins, as a potential avenue for treatment

1.3.1. The interconnected pathways resulting in Alzheimer's disease

It has become abundantly clear, due in no small part to the failure of several drug candidates aimed at inhibiting the amyloid pathway, that targeting the amyloid cascade alone is insufficient in curbing the cognitive decline associated with AD (Hawkes, 2017). This is unsurprising given the complex and interconnected nature of the many facets of the disease. The failures of drug candidates targeting the amyloid pathway do not necessarily rule out Aβ as a major player however, instead these failures suggest an emphasis would be better placed on the points of intersection between the various involved pathways rather than simply targeting one pathology at a time. This is further evidenced by recent genome-wide association studies that have shown risk factor genes for AD that encode for proteins involved in lipid homeostasis, regulation of microglial activation as well as several genes in mitochondrial DNA (Bin *et al*., 2019, dos Santos *et al*., 2017, Kunkle *et al*., 2019). This inflammatory response in the brain is mediated by two cell types, astrocytes and microglia. When the microglia become activated into either M1 or M2 forms, typically through injury, M1 microglia produce pro-inflammatory cytokines while M2 microglia have an anti-inflammatory response (Yang *et al*., 2019). It has been demonstrated in mice that activated microglia and astrocytes further promote Aβ plaque formation (Guo *et al*., 2002) which can compound the initial inflammatory signal as Aβ induces further activation of the complement system, which further facilitates microglial activity, resulting in a colocalization of microglia with amyloid plaques (Matsuo *et al*., 2017). The aggregation of Aβ can further enhance an inflammatory response through enhancing the Nuclear Factor κ-light-chain-enhancer of activated B cells (NF-κβ) and Mitogen-activated protein kinase (MAPK) pathways which have both been shown to stimulate expression of pro-inflammatory genes (Lim *et al*., 2013, He *et al*., 2020, Kim *et al*., 2018). The NF-κβ pathway further increasing BACE1 gene expression as well as BACE1 activity resulting in further Aβ

production. This has further ramifications as Aβ results in increased activity of the nicotinamide adenine dinucleotide phosphate (NADPH) oxidase complex found within microglia which, combined with the M1 microglia's natural production of inducible nitric oxide synthase (iNOS), results in an increase in oxidative stress (Neniskyte *et al*., 2016).

ApoE has also been shown to interact with both Aβ and tau, with this interaction being facilitated by lipid rafts. Further, it has been seen to influence the removal of Aβ from the brain. While ApoE is the major AD-related protein involved in lipid homeostasis, a non-protein molecule of particular interest to AD is cholesterol. Cholesterol itself is present in myelin as well as neuronal and glial membranes and plays a major role in the formation of lipid rafts, lipid rafts being where ApoE interacts directly with Aβ and tau (Rushworth *et al*., 2013). Decreased levels of cholesterol have been implicated in an uptick of Aβ production however elevated levels have also been implicated in early stage AD indicating that any divergence from cholesterol homeostasis can increase AD susceptibility (Merched *et al*., 2000, Ghribi *et al*., 2006). Lipid rafts are also where Aβ can form complexes with gangliosides (Grimm *et al*., 2012). These complexes have been seen to disrupt cell membrane stability and permeability, which has been seen to confer Aβ cytotoxicity and is also regulated by cholesterol levels. Cholesterol also induces colocalization of BACE1 and APP at lipid rafts leading to increased Aβ production (Vetrivel *et al*., 2009b). Increased APP cleavage leads to increased extracellular cholesterol levels resulting in further disruption of homeostasis. Oligomeric tau is also detrimental to the brain due to both its ability to disrupt membranes resulting in toxicity as well as disrupt autophagy of Aβ aggregates (Ait-Bouziad *et al*., 2017). Further ways in which tau processing proteins are intertwined with Aβ is through the tau truncating protein caspase 3, which catalyses tau truncation at its C-terminus and exhibits higher activity in the presence of Aβ (Cotman *et al*., 2005).

1.3.2. *Potential viability of a multitarget approach for treating Alzheimer's disease*

Given the numerous proteins and pathways directly implicated in Alzheimer's disease it has become abundantly clear that no single protein is solely responsible for the onset of AD. As a result, once symptoms begin to manifest in patients, treatments that target single proteins are unlikely to have permanent affects, even evidenced by the recent success of Biogen's aducanumab anti-Aβ antibody treatment (Biogen, 2019). This treatment indicates that continual removal of Aβ plaques may be possible and effective, however given that aducanumab treatment requires repeated intravenous administration, it clearly indicates that the underlying

disruption in brain homeostasis remains. This is likely due to the interconnectedness of all the implicated pathways in AD. As a treatment corrects one proteins levels, Aβ levels in the case of aducanumab, the remaining disrupted pathways restore the levels over time. Targeting of multiple pathways in tandem may thus increase the time it takes for these pathways to return to the disrupted levels. This lends value to the exploration of further treatments targeting BACE1 despite the recent clinical failures.

1.4 The BACE1 enzyme as a target

1.4.1.   BACE1 localization and function

The BACE1 protein is present in the majority of tissues in the body however it is expressed to a larger extent in the brain, where it is predominantly expressed in neurons. It has however been reported that although astrocytes do express BACE1 to a lower extent, they are capable of expressing enzymatically active BACE1 (Vassar *et al*., 1999; Robner *et al*., 2004). It is a transmembrane enzyme with a pH optimum of 4.5 resulting in it being most active inside the lumen of endosomes and the Golgi apparatus where it is colocalized with APP prior to APP cleavage (Skovronsky *et al*., 2000). It is also present on the cell membrane in lower concentrations as well as in synaptic vesicles (Lundgren *et al*., 2015). It is active as a monomer however dimerization is necessary for maximal activity (Schmechel *et al*., 2004). Although it is named after its ability to cleave APP, it has been reported to also act on neuregulin-1 through which it facilitates maturation and maintenance of muscle spindles (Cheret *et al*., 2013). Other BACE1 substrates include the seizure protein 6 and seizure 6 like protein, which have been linked to synapse dysfunction (Pigoni *et al*., 2016). BACE1 has also been implicated in the regulation of astrocyte and neuron formation through cleavage of the Jagged1 signalling protein (Hu *et al*., 2013). It also plays a role in myelination of neurons as well as neuroplasticity as mentioned prior.

1.4.2.   BACE1 domain structure and notable features

The BACE1 enzyme is an aspartic protease synthesized as a 501 amino acid peptide with five domains, which include the C-terminal cytosolic domain, the transmembrane domain and the catalytic domain. There is also a pro peptide region and an N-terminal signal peptide, both of which are cleaved off resulting in the fully active enzyme (Dislich and Lichtenthaler, 2012). The transmembrane and cytosolic domains play a role in BACE1 regulation as both the transmembrane and cytosolic domain assist in dimerization whilst the cytosolic domain is

required for trafficking of the protein (Yan *et al*., 2001; Schmechel *et al*., 2004). The BACE1 protein has three disulphide bonds, with each being required for activity as they maintain the correct conformation of the active site cleft. It is also the subject of extensive post-translational modifications. The protein is glycosylated four times, with these glycosylations not being integral to activity, however they are required for maximal activity (Huse *et al*., 2000). It also contains four palmitoylation sites as well as seven acetylation sites. Palmitoylation occurs on the cytosolic domain and assists in the formation of homodimers (Vetrivel *et al*., 2009a). Acetylation is not permanent and only occurs in immature BACE1 forms to facilitate transport (Costantini *et al*., 2007).

1.4.3. The catalytic domain

The catalytic domain of the enzyme has two catalytically active aspartate residues, characteristic of aspartic proteases and consists of an N-terminal and C-terminal lobe. One of the major features of the catalytic domain is the presence of an antiparallel hairpin loop commonly called a "flap", which is a common feature of aspartic proteases. This flap is integral for BACE1 inhibitor design as it plays a major role in allowing substrates access to the active site cleft. It is also implicated in ensuring the correct conformation of substrates in the binding site as well as the removal of the products of hydrolysis from the active site cleft (Xu *et al*., 2012). The large binding site of BACE1 in some cases, can accommodate up to 11 amino acids (Turner *et al*., 2001). The three-dimensional (3D) BACE1 structure is represented in Figure 1.2.

1.4.4. Catalytic site specificity

The most commonly used assays BACE1 activity assays utilize the so-called "Swedish mutation" APP (Mancini *et al*., 2011). This mutant is the result of a double mutation of the wild-type APP which is correlated with an increased susceptibility to AD due to the mutation increasing the affinity of the substrate for BACE1 cleavage (Haass *et al*., 1995). This substrate coupled with various combinations of fluorophores and quenchers is used for its high sensitivity for BACE1 proteolytic activity as well as its applicability to the multi-well format. The sensitivity of these assays is a major concern for BACE1 assays however, as the BACE1 catalytic sites affinity against the β-site of APP is low, even with the Swedish mutation (Turner *et al*., 2001). This has led to the development of optimized substrates for the BACE1 catalytic site, with the optimal amino acid cleavage site containing a leucine at P1, valine at P3 and

alanine or valine at P2'. BACE1 diplays higher affinity for hydrophobic residues at P1 and P3 (Tomasselli *et al*., 2003).
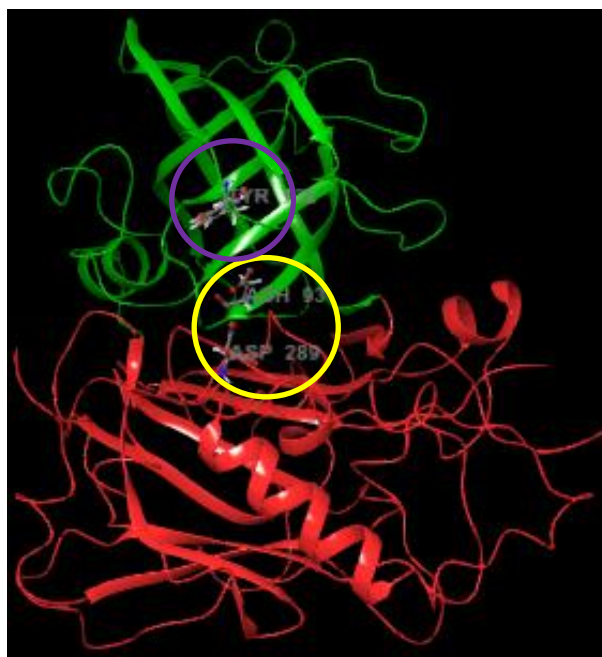
Figure 1.2: The 3D BACE1 catalytic domain. The three-dimensional ribbon representation of the BACE1 catalytic site (Protein DataBank (PDB) ID 5HU1, visualized using the Schrodinger Biologics suite (2018-2) with the tyrosine 71 residue (Encapsulated in purple) and catalytic aspartate dyad represented in grey (Encapsulated in yellow). N-terminal lobe represented in green, C-terminal lobe represented in red.

1.4.5. BACE1 proteolytic mechanism

The BACE1 protein functions similarly to other aspartyl proteases, utilizing the aspartate dyad in conjunction with a preserved water molecule in its active site to bring about peptide bond cleavage. Initially, the Asp32 residue is protonated while the Asp228 residue is not. As seen in panel A of Figure1.3 the Asp32 residue functions as a proton donor, passing on a hydrogen atom to the peptide bond's carbonyl oxygen. Simultaneously, the unprotonated Asp228 takes on a proton from the active site water molecule, creating a nucleophilic hydroxyl group which then acts on the peptide bond of the substrate, forming a geminal diol intermediate. This is the rate-limiting step of the reaction. In panel B, cleavage of the peptide bond occurs when Asp32 sequesters a proton from the intermediate, while Asp228 donates a proton to the nitrogen atom of the peptide bond. Finally in panel C, the N-terminus dissociates from the Asp32 residue resulting in the two cleavage products (Paul *et al*., 2016).
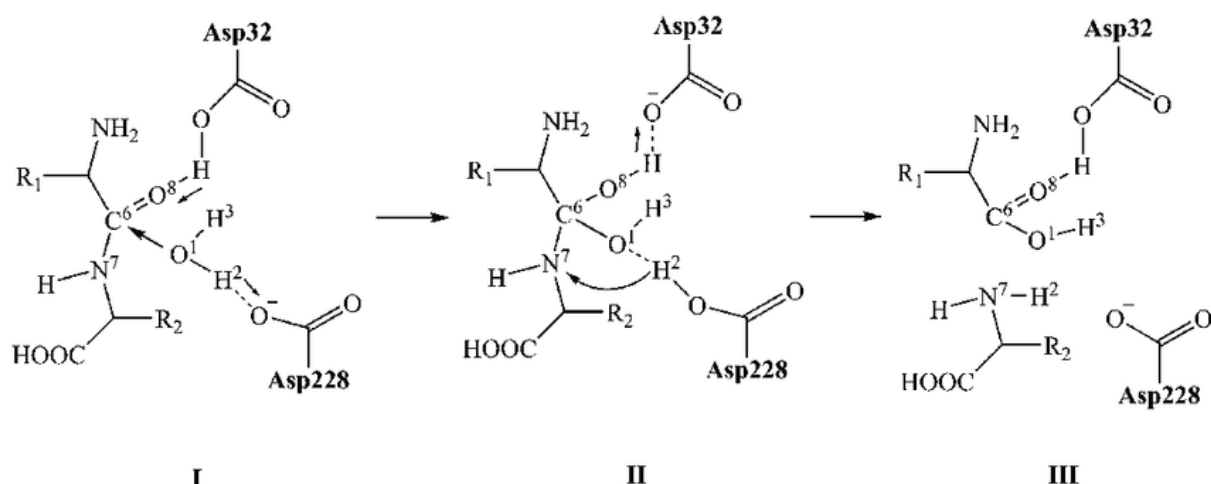
Figure 1.3: The proposed three-step cleavage of a peptide bond by the aspartate dyad of BACE1. The catalytic aspartates are represented on the left, with the substrate peptide bond on the right. The catalytic water molecule is also displayed. Arrows indicate atom interaction. Solid lines indicate covalent bonds, dashed lines indicate hydrogen bonds. Taken from Paul *et al*., 2016.

1.4.6. Concerns regarding BACE1 as a target for Alzheimer's disease treatment

A concern for the use of BACE1 inhibition as a form of AD treatment stems from the BACE1 homolog, β-site APP cleaving enzyme-2 (BACE2). The BACE2 amino acid sequence possesses 64% similarity, with 51% identity, to that of BACE1. This similarity in structure is highlighted in Figure 1.4 in which the structures of BACE1 and BACE2 are superimposed on each other. BACE2 is present in most tissues of the human body, much like BACE1, and is expressed highest in the pancreas. It is not highly expressed in the brain however, though BACE2 activity has been reported in the brain. Studies have been carried out on BACE1 and BACE2 knockout mice where BACE2 knockout mice displayed a healthy phenotype compared to the complicated phenotypes of BACE1 knockout mice summarized in section 1.2.4. Mice incapable of expressing either BACE1 or BACE2 presented higher lethality, resulting in a survival rate of 40%, whereas mice incapable of expressing either BACE1 or BACE2 exhibited no such lethality (Dominguez *et al*., 2005). This is of concern as any compound that can inhibit both BACE1 and BACE2 due to the high degrees of similarity may result in detrimental side-effects. With regards to its ability to cleave APP, BACE2 can cleave at the APP β-cleavage site, however it favours cleavage at a site closer to the α-cleavage site (Yan *et al*., 2001b).
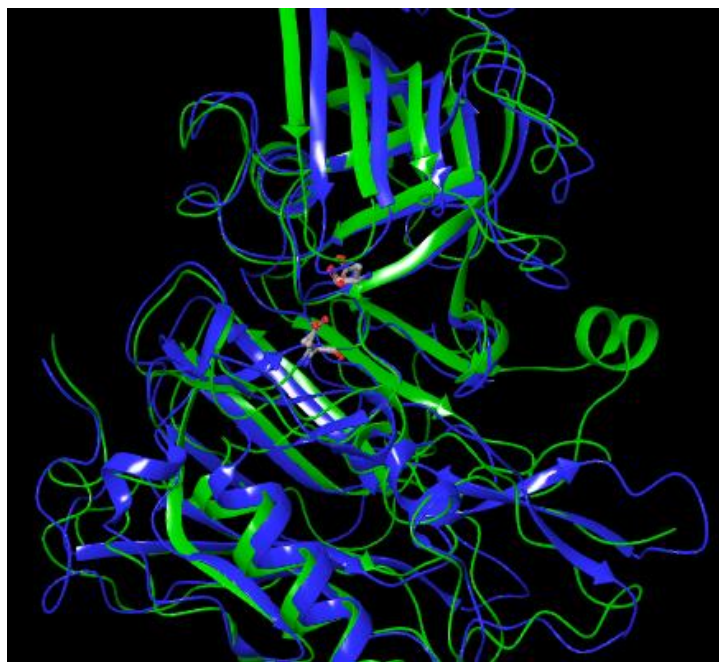
Figure 1.4: BACE2 structure superimposed onto BACE1. The BACE1 (5HU1) and BACE2 (2EWY) structures were superimposed using the Schrodinger Biologics suite (2018-2) aligning their respective catalytic aspartates. BACE1 is coloured in blue while BACE2 is represented in green. The catalytic aspartate dyad is represented in grey.

1.4.7. Generation of BACE1 protein for study

A protocol for reliable production of soluble, active BACE1 was published in 2003 (Sardana *et al.*, 2003). This bacterial expression however resulted in 3.5 mg/L of culture due to aggregation during affinity purification. Large scale, high yield expression of BACE1 for study has still proven to be difficult due to the protein's expression in inclusion bodies as well as its tendency to aggregate. Expression in mammalian cells, such as human embryonic kidney (HEK)-293 and Chinese hamster ovary (CHO)-K1 cell lines has also been carried out (Emmons *et al.*, 2008), however since the protein has a total of four glycosylation sites, BACE1 generated in mammalian cell lines is difficult to use for crystallographic studies as these glycosylations often lead to low resolution crystals (Chang *et al.*, 2007). Further, mammalian cell expression is associated with low protein yields, relative to bacterial expression with BACE1 expression in HEK-293 cells yielding 1.5 mg/L and expression in CHO-K1 cells yielding 0.75 mg/L (Emmons *et al.*, 2008). Bacterial expression has drawbacks however, as the protein has three disulphide bonds which are required for activity, necessitating further steps to ensure the protein is correctly folded (Tomasselli *et al.*, 2008). Furthermore the glycosylations have been demonstrated to be necessary for maximum proteolytic activity (Sathya *et al.*, 2012). The yields in *Escherichia coli* (*E. coli*) are however six times higher than those found in mammalian cells

with the protocol yielding the highest amount of BACE1 producing 7 mg/L of recombinant protein while also producing high resolution crystal structures (Tomasselli *et al.*, 2008). This expression however requires a lengthy three-day incubation to allow for refolding of the recombinant protein.

1.5 Drug Discovery

The discovery of drugs is a process that spans around 15 years from initial target identification, as seen in Figure 1.5. The major phases of the drug discovery pipeline include target identification, development and application of screening assays, preclinical assessments of safety, usually carried out in animal models, clinical trials, ending finally in approval for sale (Roses, 2006). After the identification of a target and confirmation of its drugability, the screening process can be carried out in two ways, random high-throughput screening or the rational drug design approach. The random approach relies on the screening of large numbers of drug-like compounds against the protein target whereas the drug design approach may be further divided into structure-based and ligand-based. Structure-based techniques utilize structures of the target protein derived from nuclear magnetic resonance spectroscopy (NMR), crystallography or homology modelling to design ligands specifically for the active site. These drug design techniques include de novo drug design as well as fragment-based approaches (Murray and Rees, 2009; Congreve *et al.*, 2007). Ligand-based methods utilize pharmacophore models which compiles common features and functional groups found in known ligands of the target protein and then allow screening for compounds that possess similar structures. Another ligand-based approach utilizes the technique quantitative structure activity relationship (QSAR) as a way of designing and optimizing known ligands (Kubinyi, 1997).
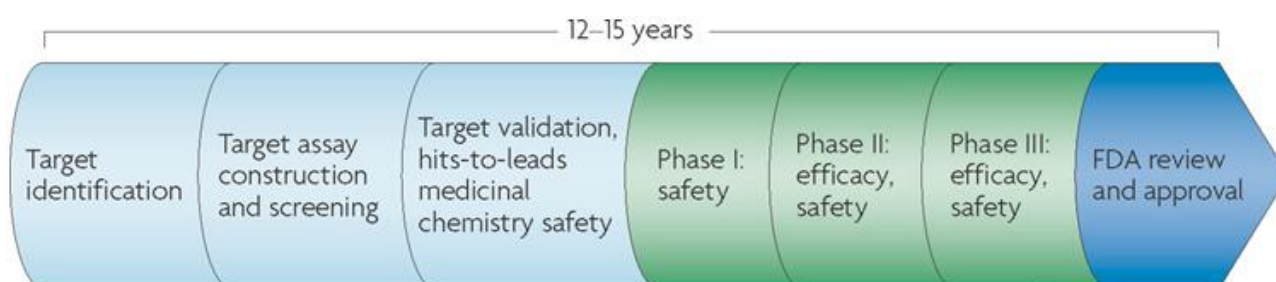
Figure 1.5: The drug discovery pipeline from initial target identification to clinical trials and FDA approval. The initial drug development stages including target identification, compound screening and hit identification are indicated in light blue, clincal trials indicated in green with the end result of FDA approval indicated in dark blue. Taken from Roses, 2006.

1.5.1 *Differential Scanning Fluorimetry*

Differential scanning fluorimetry (DSF) is a ligand binding assay that uses the random-based approach to screen for binding molecules. It takes advantage of the process of denaturation of proteins when exposed to high temperatures as well as the ability of bound ligands to increase protein stability. This is carried out through the use of a fluorescent dye capable of binding to hydrophobic protein residues. Over the course of a DSF experiment, the temperature of the test solution is gradually increased. As this occurs, the protein in question begins to denature exposing its hydrophobic residues, allowing the fluorescent dye to bind. This binding is detected by an increase in fluorescence until a temperature is reached where the protein begins to dissociate from the dye. The relationship between temperature and fluorescence is then used to determine the proteins melting temperature (Tm) as shown in Figure 1.6. When in the presence of a ligand, this Tm increases as the ligand enhances the stability of the bound protein. This difference is referred to as a thermal shift (Niesen *et al*., 2007). DSF is highly amenable to automation and, through the use of multi-well plates allows for high-throughput screening (Menzen and Friess, 2012). It does allow for the elucidation of the proteins standard Gibbs energy of unfolding ($\Delta_u G^\circ$) which can be used as an inference of the proteins thermal stability (Wright *et al*., 2017). The DSF assay also allows the elucidation of the class of unfolding, whether single-state or two-state unfolding.

The DSF assay does have drawbacks however as it only detects molecules capable of binding to the target protein, this does not always correlate to inhibition however. This drawback may potentially be why the use of the DSF assay to screen for unknown BACE1 inhibitors has not been reported, with the only reported use of DSF being carried out on BACE1 being a study assessing the viability of the DSF assay utilizing known BACE1 binding compounds (Lo *et al*., 2004). With the recent increase in the use of proteolysis targeting chimers (PROTACs) as a means to induce protein degradation however, the downside of being unable to detect inhibition is irrelevant. Furthermore, given the failure of the Verubecestat in phase 3 trials despite the compounds potent BACE1 inhibitory ability, new means of targeting and reducing BACE1 levels, such as through the use of PROTACs, may be necessary.
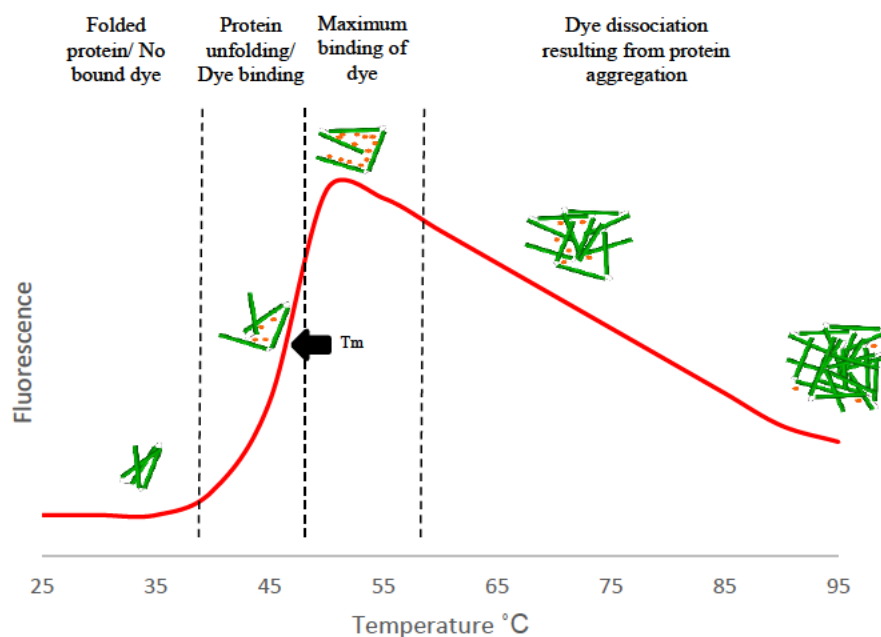
Figure 1.6: A standard melt curve of a protein exposed to a differential scanning fluorimetry assay. The stages of protein denaturation and dye binding are demonstrated. The protein melting temperature (Tm) is indicated.

1.5.2 Cellular Thermal Shift Assay

Another method that utilizes the same principle as DSF is the cellular thermal shift assay (CETSA). Instead of determining the Tm of the pure protein, it measures the T_{agg} of the protein in the cellular environment (Molina *et al.*, 2013). This assay is carried out by growing cells overexpressing the protein of interest, lysing said cells and subsequently exposing fractions of that lysate to different temperatures. The stability of the protein is then determined by isolating the soluble fraction of proteins after heating and carrying out a western blot using antibodies against said protein with the presence of the protein in the soluble fraction being taken as confirmation that the protein has yet not melted. The temperature at which 50% of the protein is no longer present in the soluble fraction is referred to as the T_{agg} (Martinez *et al.*, 2018). This assay has the benefit of testing if ligands can not only bind to the target protein in a cellular environment, but also if they are capable of entering the cell to interact with the protein. It also has the benefit of not relying on purified BACE1 as is the case with DSF and the commonly used fluorescence resonance energy transfer (FRET), which is of value given the reported difficulties with BACE1 purification. The disadvantage of this technique however is that it is not as scalable for screening of large numbers of compounds as compared to the previously mentioned DSF and FRET assays.

The use of the CETSA for the study of BACE1 has not been reported however and thus could be a novel means of detecting intracellular BACE1 binding compounds. Of further interest with regards to the screening of BACE1 binding compounds is that potentially insight into the intracellular binding of both the BACE1 monomer and dimer could be obtained given BACE1 has been reported to present natively as a homodimer (Westmeyer *et al*., 2004). It is not possible in assays such as DSF and FRET to easily separate BACE1 dimers from monomers, however since the CETSA analysis is carried out*via*western blot, the dimer is separated from the monomer naturally through electrophoresis. This ability to separate dimers could potentially allow for the screening of compounds that target the BACE1 dimer specifically.

1.5.3 Molecular Docking

Molecular docking is a structure-based drug design technique for in silico screening. It is commonly used for screening for BACE1 ligands due to the ubiquity of BACE1 crystal structures available on the Protein DataBank (PDB). Compounds are screened using software that predicts binding affinity of the ligand in the active site of the protein in question. This affinity is assigned a score which can then be compared to the scores of all other compounds screened in this way. Initial models of BACE1 are typically generated using a 3D structure of BACE1 co-crystallized with a ligand, imported from the PDB. The position of the ligand is then used to define a receptor grid. Compounds that are rendered in 3D can then be screened against this receptor grid. This method is effective for initial hit identification as it allows for rapid and inexpensive screening of large compound libraries. A commonly used molecular docking program is the Glide (Grid-based Ligand Docking with Energetics) program. Initially, a compound is separated into several conformations, generated through rotations around the compounds rotatable bonds (Friesner *et al*., 2004). These conformations are then fit into the receptor grid active site, with their positions being assessed with regards to location inside the receptor grid. Conformations that fit correctly into the receptor grid, then have their energies minimized using Optimized Potential for Liquid Simulations-All Atoms (OPLS-AA) forcefields. The lowest energy conformations are then subjected to Monte-Carlo simulations to determine the final, lowest energy conformations in the receptor grid. Lastly interactions between the receptor and the ligand are assessed and assigned scores, referred to as Glide scores, where lower the Glide scores indicate tighter binding of the ligand to the receptor.

Ligand conformations will then typically be exposed to molecular dynamics simulations to validate and better define the bound ligand conformation (Kumar *et al*., 2017).

1.5.4   *The zebrafish embryo toxicity test*

The zebrafish, *Danio rerio*, is a bony fish native to regions of India and its surrounding countries. Zebrafish were first used in the 1970s as models to not only measure toxicity, but studies were also carried out on their neuronal structure as well as on their spawning cycle and clutch sizes (Eaton and Farley, 1975, Eaton and Farley, 1974, Thomas, 1975). These studies resulted in the genome of the zebrafish ultimately being fully sequenced in 2001 by the Wellcome Trust Sanger institute (Howe *et al*., 2013). The use of zebrafish as not only a toxicity model but a model organism has increased substantially since then due to the several benefits over other, more established models (Teame *et al*., 2019). The major benefits include the frequent production of large numbers of embryos that are largely transparent throughout the developmental stages. These benefits make zebrafish embryo toxicity test (ZFET) amenable to high throughput screening due to the large and frequent number of embryos, with toxicity endpoints being assessable in real time due to their transparency. Furthermore, studies have demonstrated that zebrafish have orthologues for around 70% of all genes found in humans, contributing to their viability as a model organism for human diseases.

The ZFET was established as a means to quantify toxicity using zebrafish embryos. This assay was adopted by the Organisation for Economic Co-operation and Development (OECD) with strict guidelines being created for its use, to ensure reliability (OECD, 2013). These guidelines layout the specific endpoints to take note of when carrying out the ZFET, stipulating that the determination of lethality be carried out as a function of four observations. The first of these observations being the coagulation of an embryo, resulting in a non-viable embryo incapable of developing further. Secondly, the lack of somite formation along the tail of the fish indicates the presence of toxicity, typically resulting in deformities in the tail of the fish as well as observable spinal curvature. The third toxicity endpoint is the non-detachment of the tail from the main body of the embryo which is observable after 24 hours of development and finally the lack of any heartbeat from 48 hours onward confirms lethality. All of these endpoints are easily identifiable over the course of the assay, allowing for simple, high throughput determination of compound toxicity.

Teratogenicity is another property that is measurable using a modification of the ZFET, making use of the transparent nature of the embryos even after being fully developed. Zebrafish grown over the course of five days are typically possess all of the structures present in adult zebrafish allowing for the assessment of how correctly these structures are formed. This scoring system, first established by Panzica-Kelly and colleagues provided a reliable means with which to quantify levels of deformity (Panzica-Kelly *et al*., 2010). This method relies on the assessment of nine observable structures in the five-day old embryo, with lower scores denoting a greater degree of malformation, as displayed in Figure 1.7. This method allowed for the calculation of a teratogenicity index (TI) through the determination of the no-observed-adverse-effects level, which is denoted as the concentration at which no change in the morphological score was observed. This method has since been updated, rather than relying on the determination of the no-observed-adverse-effects level, more easily calculated values are now used. These values being the concentration at which lethality of a compound is 50% (LC_{50}) and the concentration at which the morphology score is 50% of the maximum score (EC_{50}) (Alafiatayo *et al*., 2019). The ratio between the LC_{50} and the EC_{50} being the TI, with teratogenicity being denoted if the TI is greater than 1. This allows for the zebrafish to function as a robust means to measure lethality and teratogenicity, making it a useful model in the drug development pipeline.
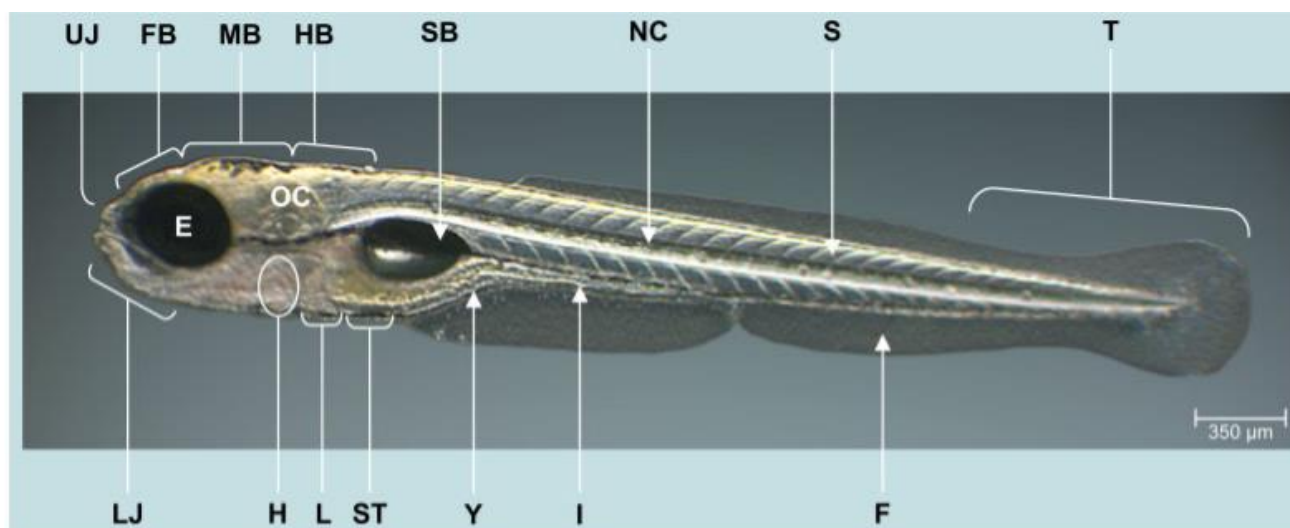
Figure 1.7: The profile of a five-day old zebrafish, denoting each of the major structures. The upper jaw denoted by UJ, The forebrain denoted by FB, MB denotes the midbrain, HB, the hindbrain, OC, the otic capsule, E, the eye, SB, the swim bladder, NC, the notochord, S, the somites, T, the tail, LJ, the lower jaw, H, the heart, L, the liver, ST, the stomach, Y, the yolk, I, the intestine and F denoting the fin. Taken from Panzica-Kelly, *et al*. 2010.

1.6 Hypothesis, Aims and Objectives

The cellular thermal shift assay is a powerful tool for detecting intracellular target engagement of protein ligands, however it has yet to be utilized in the search for BACE1 binding ligands. It was hypothesized that the CETSA could be used to aid in the search for BACE1 ligands. The aim of this study was to assess the thermostability of BACE1 and determine the viability of using thermostability as a means to detect BACE1-ligand binding in light of its reportedly high thermostability. This was carried out through the following objectives:

1. Recombinantly expressed BACE1 yields, greater than those found in literature, were determined to be possible.

2. The value of utilizing thermostability of BACE1, for compound screening, was assessed and confirmed.

3. The comparability between *in vitro* and in silico BACE1 ligand screening was explored.

4. Intracellular binding capabilities of BACE1 ligands, determined in Objective 2, were evaluated.

5. The use of the compound from Objective 4, that was detected to bind BACE1 intracellularly, for use as a precursor to clinically relevant BACE1 ligands was examined.

6. The potential of the compounds from Objective 2 and 5, for use in *in vivo* tests were assessed.

CHAPTER 2: MATERIALS AND METHODS

2.1 Plasmid preparation and *Escherichia coli* transformation

2.1.1 *rBACE1-pGEX4T-1 plasmid synthesis and the transformation of E. coli*

The gene sequence for human BACE1 was obtained from GenBank (GeneID: 23621). The amino acid sequence for human BACE1 was obtained from UniProt (BACE1_HUMAN). The peptide sequence was assessed with regards to hydrophobicity using the online protein analysis tool, ProtScale (Swiss Institute of Bioinformatics, Switzerland), as well as analysed for cleavage sites from common proteases using PeptideCutter (Swiss Institute of Bioinformatics, Switzerland). The nucleotide sequence was analysed for restriction digestion sites using the online tool NEBCutter (New England Biolabs, USA). The sequence coding for the truncated BACE1$_{43-454}$ gene was then synthesised into the pGEX4T-1 plasmid by GenScript (USA). This construct is henceforth referred to as the BACE1$_{43-454}$–glutathione-S-transferase (GST) plasmid. This gene was subjected to codon optimization to ensure maximal expression in *E. coli* cells (Burgess-Brown *et al*., 2008). *E. coli* BL21 (DE3) cells were transformed using calcium chloride transformation as per Dagert and Ehrlich (1979). Briefly, a glycerol stock of untransformed *E. coli* BL21 (DE3), purchased from New England Biolabs (USA), was plated onto 2xYT agar (16 mg/ml tryptone, 10 mg/ml yeast extract, 5 mg/ml NaCl, 100 µg/ml ampicillin, 1.5% (w/v) bacteriological agar, all purchased from MilliporeSigma, USA) and incubated overnight at 37°C. From this a single colony was selected and grown in 2xYT broth (16 mg/ml tryptone, 10 mg/ml yeast extract, 5 mg/ml NaCl, 100 µg/ml ampicillin) for 16 h at 37°C, shaking at 200 revolutions per min (RPM). This culture was then diluted 1:100 with fresh 2xYT broth and grown in the same conditions until an optical density (OD), at a wavelength of 600 nm, of 0.3 - 0.4 was reached. The culture was then transferred to ice-cold centrifuge tubes and centrifuged at 4000 *xg* for 10 mins at 4°C. The pellet was subsequently resuspended in ice-cold sterile CaCl$_2$ solution (60 mM CaCl$_2$, 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer, pH 7.0, all purchased from MilliporeSigma, USA). The resuspended pellet was then centrifuged under the same conditions as before and the pellet resuspended in a volume of CaCl$_2$ solution equal to 1 / 50th the final culture volume.

The resulting cell mixture was then incubated with 1 / 20th the volume of BACE1$_{43-454}$ –GST plasmid for 30 mins on ice. These cells were then heat-shocked at 42°C for 90 s and further incubated on ice for two mins. The cell solution was then added 1:4 to Super Optimal broth

with Catabolite repression (SOC) media (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose, all purchased from MilliporeSigma, USA) and incubated at 37°C for 1 hour with gentle agitation. The resulting mixture was then plated onto a pre-warmed 2xYT plate and incubated for 16 h, again at 37°C. This plate was then used to generate glycerol stocks of the transformed *E. coli* cells. Herein, ten 2 ml, 2xYT aliquots containing ampicillin (MillporeSigma, USA) were inoculated with single colonies, with these cultures being incubated at 37°C for 16 h with 200 RPM shaking. After this sterile, glycerol was added to a final concentration of 7.5%. These glycerol stocks were then stored at -70°C.

2.1.2 *Plasmid Isolation and restriction digestion*
The glycerol stock containing the *E. coli* BL21 (DE3) cell line (New England Biolabs, USA) were plated onto 2xYT agar and incubated overnight at 37°C. A single colony was then grown in 2xYT broth overnight as before. From this overnight culture, plasmid DNA was isolated using the GeneJet plasmid DNA isolation kit (Thermo Fisher Scientific, USA) as per the instructions provided by the manufacturer. Restriction digestion was carried out on the isolated DNA using EcoR1 (1U) and BamH1 (2U) in 2 x "Tango" buffer (Thermo Fisher Scientific, USA) overnight at room temperature.

2.1.3 *Agarose gel electrophoresis*
The samples that were exposed to restriction digestion, were analysed on a 1% (w/v) agarose gel alongside the undigested plasmid, as well as 5 µl of O'GeneRuler 1 kb DNA ladder (Thermo Fisher Scientific, USA). This agarose gel was made by dissolving agarose to a final concentration of 1% (w/v) in 1 x Tris (hydroxymethyl) aminomethane (Tris)-acetate-ethylene diamine tetra-acetic acid (EDTA) buffer (TAE buffer) (40 mM Tris-HCl, 20 mM glacial acetic acid, 1 mM EDTA) containing 0.5 µg/ml ethidium bromide (Invitrogen, USA). Visualization was carried out using a GeneSys G-box gel visualization system (Syngene, India) with ultraviolet light.

2.1.4 *Plasmid sequencing and purity analysis*
To assess the final concentration and purity of the isolated plasmid, A_{280} and A_{260} values were measured using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA). This isolated plasmid DNA was sent to Inqaba Biotec (RSA) for DNA sequencing.

2.2 Expression of rBACE1 in *Escherichia coli*

2.2.1 Determining the solubility of rBACE1-GST

2xYT broth with ampicillin was inoculated with a single colony of transformed *E. coli*. This was then grown overnight at 37°C while shaking at 200 RPM. This overnight culture was then used to inoculate two separate 2xYT broths, at a 1 in 100 dilution. The resulting cultures were then incubated at 37°C, 200 RPM until an OD_{600} of 0.4 was reached. After this, 1 mM IPTG (Thermo Fisher Scientific, USA) was added and the resulting induced culture incubated at 37°C, 200 RPM for 4 h. After induction was complete, the cells were harvested through centrifugation at 4000 *xg* for 10 mins at 4°C.

The supernatants were discarded and the pellet resuspended in phosphate buffered saline (PBS) pH 7.4 (2.7 mM KCl, 1.8 mM KH_2PO_4, 10 mM Na_2HPO_4, 150 mM NaCl in 500 ml), equal in volume to 10% of the original culture volume. The resuspended sample was then sonicated 5 times for one min with one min intervals on ice and subsequently centrifuged at 10 000 *xg* for 20 mins at 4°C. The supernatant was stored at 4°C as the PBS soluble sample. The pellet was then resuspended in resolubilization buffer (50 mM Tris-HCl, 8 M urea, 10 mM 2-mercaptoethanol, pH 9.0, all purchased from MilliporeSigma, USA) equal in volume to 5% of the original culture volume.

2.2.2 Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis and Western Blot Analyses

After expression, sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis was carried out on all samples under non-reducing conditions in accordance with Laemmli (1970). Briefly, a 1 mm thick, polyacrylamide gel, comprised of a running gel and stacking gel containing 12.5% and 4% acrylamide respectively, was loaded with a molecular weight marker alongside protein samples treated with non-reducing treatment buffer (125 mM Tris-HCl, 4% SDS, 20% glycerol, with bromophenol blue). Three molecular weight markers were used throughout the study with the three being: the SeeBlue Plus2 Pre-stained Protein Standard molecular weight marker, Spectra™ Multicolor Broad Range Protein Ladder molecular weight marker and Pageruler Plus prestained protein standard, purchased from Thermo Fisher Scientific (USA). Each gel was electrophoresed at 20 mA in tank buffer (250 mM Tris-HCl, 192 mM glycine, 0.1% (w/v) SDS) to completion. Gels were then either stained or analysed *via* a western blot. Gels were stained using Coomassie blue staining solution

(0.125% (w/v) Coomassie brilliant blue R-250 (MilliporeSigma, USA), 50% (v/v) methanol (RadChem, USA), 10% (v/v) glacial acetic acid (RadChem, USA)). Molecular weight values for bands corresponding to unknown proteins were determined using molecular weight markers. The distance travelled by each marker was plotted against each bands corresponding log(molecular weight). A line of best fit was derived and used to calculate the molecular weight of any bands representing proteins of unknown molecular weight.

Western blotting was carried out by first transferring proteins from a polyacrylamide gel to a nitrocellulose (MilliporeSigma, USA) or a polyvinylidene difluoride (PVDF) membrane (Pall Corporation, USA), with the PVDF membrane being pre-wet in methanol prior to transfer. Transfers were carried out using a Bio-Rad overnight transfer system (Bio-Rad, USA). This was carried out overnight in blotting buffer (25 mM Tris-HCl, 190 mM glycine, 20% (v/v) methanol, 0.1% (m/v) SDS). After protein transfer, the membrane was blocked for an hour using 5% (w/v) low-fat dry milk powder dissolved in Tris buffered saline (TBS; 20 mM Tris-HCl, 200 mM NaCl, pH 7.4) containing 0.1% Tween-20. After blocking, the blot was washed three times with TBS-Tween after which, primary antibodies in TBS-Tween, containing 0.5% Bovine serum albumin fraction V (BSA, Biowest, France) were added. The blot was then incubated at room temperature for three h with gentle mixing. Both the rabbit anti-BACE1 primary IgG antibodies (MillporeSigma, USA) and chicken anti-GST IgY antibodies (kindly provided by Professor Theresa Coetzer) were used at a concentration of 2 μg/ml. The blot was again washed with TBS-Tween three times with secondary antibody then being added. The goat anti-chicken (Advansta, USA), as well as the goat anti-rabbit IgG antibodies (MilliporeSigma, USA) were used at dilutions of 1:10 000, both of which were conjugated to horse radish peroxidase (HRPO). Secondary antibody incubation was carried out for one hour at room temperature, again with gentle mixing. The blot was then washed and the protein bands visualized with an enhanced chemiluminescent (ECL, Bio-Rad, USA) substrate kit and imaged using a Syngene Chemi XR5 G-box with GeneSys gel visualization software (Syngene, India).

2.2.3   Optimization of expression conditions

After confirmation of the presence of rBACE1-GST in the cell lysate, the expression conditions were optimized with regards to type of media, IPTG concentration, induction time and induction temperature. The types of media compared were 2xYT media and lysogeny broth (LB) (8 mg/ml tryptone, 5 mg/ml yeast extract, 5 mg/ml NaCl, 100 μg/ml ampicillin). The two

IPTG concentrations 0.5 and 1 mM, two different induction times, namely four h and overnight induction as well as the induction temperatures 16°C and 37°C were also compared.

2.2.4 *Use of anti-aggregation to generate soluble rBACE1-GST*

Several conditions were tested including 1% polyethylene glycol (PEG) 6 000, 10% glycerol, 1% Triton X-100, 200 mM L-Arginine and 200 mM sucrose. These were added to the cell pellet after resuspension in PBS. This was followed by sonication and centrifugation as before, with the soluble and insoluble fractions being compared on an SDS-PAGE gel.

2.2.5 *Resolubilization and refolding of rBACE1-GST*

Initially a refolding strategy was attempted using a ratio of oxidized and reduced glutathione. Briefly, after expression, the soluble and insoluble fractions of the cell lysate were separated and the insoluble fraction resuspended in resolubilization buffer equal in volume to 5% of the original culture volume. This was then diluted 1 in 20 in refolding buffer (20 mM Tris-HCl, 0.5 mM oxidized glutathione, 2.25 mM reduced glutathione, pH 9.0) with the sample then being incubated at 4°C for four h after which it was titrated to pH 8.0. Further, this was incubated at 4°C for 48 h, titrated to pH 6.8 and further incubated for another 48 h at 4°C as per. The resulting solution was then dialyzed against PEG 20 000 in regenerated cellulose, molecular weight cut-off (MWCO) 10 000, Snakeskin™ dialysis tubing (Thermo Fisher Scientific, USA).

A further technique involving the incubation of the insoluble protein fraction in a 10% N-lauryl sarcosine (sarkosyl) solution for 24 h was also tested. In this, the insoluble pellet was dissolved in PBS containing 10% sarkosyl and incubated at 4°C for 24 h, after which the solution was diluted 1 in 5 reducing the sarkosyl concentration before column purification, further, 4% Triton X-100 was then added.

2.2.6 *Glutathione-agarose column purification*

A glutathione-agarose gravity column (MilliporeSigma, USA), 1 ml in volume was swollen and equilibrated to the relevant buffer as per instructions provided by the manufacturer (Catalogue number: G4510). After this, the sample containing the GST fusion protein was incubated in the resin for one hour with constant mixing. The column was then left to settle and subsequently washed with PBS containing 1% Triton X-100 or 4% Triton X-100 (for sarkosyl solubilization) until an A_{280} reading below 0.02 was reached.

The bound protein was then eluted using elution buffer (10 mM reduced glutathione, 50 mM Tris-HCl, pH 9.0) with the fractions being collected until an A_{280} of 0.02 was again reached. The unbound fraction as well as all eluted fractions were then analysed for protein on an SDS-PAGE gel. The column was then regenerated using five column volumes of sodium borate buffer (200 mM sodium borate, 500 mM NaCl, pH 8.0), followed by five column volumes of dH_2O, then five column volumes of sodium acetate buffer (100 mM sodium acetate, 500 mM NaCl, pH 4.0) then finally being stored in storage buffer (500 mM NaCl with 1 mM NaN_3) at 4°C. This purified, uncleaved rBACE1-GST was analysed using an SDS-PAGE and western blot after which it was used for immunization of chickens for antibody generation.

2.2.7   Bradford Assay

A Bradford protein assay was then carried out on each eluted fraction in an effort to determine the concentration of the eluted protein. This assay was carried out using a NanoDrop 2000 spectrophotometer. A protein standard curve was made using a stock solution of 1 mg/ml BSA to prepare protein standard solutions ranging from 30 to 200 μg/ml, with all readings carried out in triplicate. Protein standards were mixed 1:1 with Bradford reagent (Thermo Fisher Scientific, USA). The absorbance at 595 nm was measured after incubation in the dark for 10 mins (Bradford, 1975).

2.2.8   Cleavage of GST-tag using thrombin

A comparison was made between on-column thrombin cleavage and off-column thrombin cleavage. For the on-column thrombin cleavage, the glutathione-agarose column was equilibrated with ten column volumes of PBS pH 7.4, containing 4% Triton X-100. The refolded and diluted rBACE1-GST was then added to the column and incubated for one hour at room temperature rotating on an end-over-end rotator. The column was then washed with PBS pH 7.4, containing 4% Triton X-100 with the eluent being collected in 1 ml fractions. The OD of each fraction, at 280 nm, was measured with fractions being collected until an OD of 0.02 was reached. After this the column was equilibrated with 20 column volumes of thrombin cleavage buffer (20 mM Tris-HCl, 150 mM NaCl, 2.5 mM CaCl2, pH 8.4). After equilibration the column was resuspended in 1 ml thrombin cleavage buffer containing 2 U thrombin (MilliporeSigma, USA). The resin was then incubated at 4°C for 16 h on an end-over-end rotator. After this, the column was washed with ten column volumes of thrombin cleavage buffer with 1 ml fractions of eluent being collected. The column was then washed with ten

column volumes of elution buffer to elute any remaining fusion protein as well as the bound GST. The column was then regenerated as previously described.

For the off-column thrombin cleavage, protein was eluted from the column as outlined prior. After the eluted fractions were pooled, a buffer exchange step was then carried out through the use of dialysis. The pool was placed into MWCO 10 000, Snakeskin™ dialysis tubing (Thermo Fisher Scientific, USA) and was dialyzed against 100 x volume of PBS for two h at 4°C with gentle stirring. This was followed by a further two hour dialysis against fresh PBS after which, a final overnight dialysis was carried out. After dialysis, thrombin was added to the purified protein at a final thrombin concentration of 10 units per mg of purified protein. This was then left to incubate overnight at room temperature. With the cleavage reaction being stopped with 1 mM phenylmethane sulfonyl fluoride (Thermo Fisher Scientific, USA). The thrombin was then removed through the use of a HiTrap Benzamidine FF column (MilliporeSigma, USA). The prepacked 1ml column was washed with 5 column volumes of distilled water, equilibrated with 5 column volumes binding buffer (50 mM PBS, 500 mM NaCl, pH 7.4). The protein pool was then added to the column at a steady rate of 1 ml/min. The column was washed with binding buffer with all fractions being analysed for protein content using absorbance at 280 nm. All fractions with A_{280} readings above 0.02 were pooled. The resulting protein pool was then rerun over the glutathione-agarose column (MilliporeSigma, USA) as before. This time however, the unbound fractions were collected and pooled. The resulting pool was then concentrated using a Centricon 30 000 MWCO centrifugal concentrators (Amicon, MilliporeSigma, USA; centrifuging at 4000 xg, 5 mins, 4°C).

2.2.9   *Gelatine gel zymography*

A gelatine gel zymogram was run identically to the SDS-PAGE gels described previously, with the addition of 0.1% gelatine to the running gel buffer prior to setting. After electrophoresis, the gel was washed in two changes of 2.5% Triton X-100 for 30 mins each. This was followed by incubation at 37°C overnight in sodium acetate buffer (50 mM acetate, pH 4.5). The gel was then stained using Coomassie stain as described for the SDS-PAGE gels before. Protease activity was detected by the presence of a clear band where gelatine degradation had occurred (Tam *et al*., 2004). A positive control was carried out using a recombinantly expressed BACE1 produced in HEK-293 cells (MilliporeSigma, USA).

2.3 Production of anti-BACE1 and anti-GST IgY antibodies

2.3.1 Chicken immunization with rBACE1-GST

The use of all animals for research was approved by the Biomedical Research Ethics Committee (053/15/Animal) at the University of KwaZulu-Natal. A single chicken was injected intramuscularly with 750 µl of protein-adjuvant solution, once per breast. For the first immunization, Freund's complete adjuvant (MilliporeSigma, USA) was mixed 1:1 with uncleaved rBACE1-GST solution, bringing the concentration of the rBACE1-GST protein to a final concentration of 100 µg/ml. The boosters that followed were carried out on weeks 2, 4 and 6, using Freund's incomplete adjuvant (MilliporeSigma, USA). Injections were carried out using a 2 ml syringe with a 20-gauge needle after sterilizing the surface of the breast with 70% (v/v) ethanol. Eggs were collected daily for 12 weeks, with an egg being collected on the day of first immunization as a pre-immune control.

2.3.2 Separation of IgY antibodies from chicken eggs

The separation of IgY from the proteins in chicken egg yolk was carried out as per Goldring and Coetzer, 2003. Briefly, the egg shell was carefully removed and the white discarded. The yolk sac was punctured and separated from its contents and the yolk volume was measured. The yolk was then mixed with phosphate buffer (100 mM Na_2HPO_4, 0.02% (w/v) NaN_3, pH 7.6) equal to twice the volume of the yolk. PEG 6 000 was then added to a final concentration of 3.5% (w/v). This was then centrifuged at 4 420 xg for 30 mins at 4°C. The pellet was discarded and the supernatant filtered through cotton wool. To this filtrate, additional PEG 6 000 was added to bring the total concentration up to 12% (w/v). This filtrate was then centrifuged at 10 000 xg, for 20 mins, at 4°C. The supernatant was then discarded and the pellet dissolved in phosphate buffer equal to the original yolk volume. To this, 12% PEG 6 000 was added and the solution centrifuged at 10 000 xg, for 20 mins, at 4°C again. The supernatant was again discarded and the pellet was then dissolved in phosphate buffer containing 0.1% (w/v) NaN_3. The protein concentration of this solution was then determined by measuring the A_{280}.

2.3.3 Enzyme-Linked Immunosorbent Assay

For the enzyme-linked immunosorbent assay (ELISA), each well of a 96 well microtitre plate (Corning, USA) was coated overnight at 4°C, with 150 µl of a 1 µg/ml solution of BACE1-GST in coating buffer (50 mM $NaHCO_3$, pH 9.6). The plate was then blocked with PBS-BSA for one hour at 37°C, using a total of 200 µl/well. The plate was then washed with PBS-Tween

three times. After drying, 100 µl of a 1 mg/ml solution made using the crude antibody sample collected prior, was added to each well as the 1° antibody, this was then incubated for one hour at 37°C.

The 1° antibody was discarded and the plate washed again with PBS-Tween three times followed by the addition of the 2° antibody (rabbit anti-chicken IgY conjugated to HRPO) (MilliporeSigma, USA) at a dilution of 1:10 000 (120 µl/well). This was again incubated for one hour at 37°C after which the plate was washed again using PBS-Tween. A solution of 1 mg/ml Tetramethylbenzidine (TMB, VWR Life Science Amresco Biochemicals, USA) was diluted 1:10 in citrate-phosphate buffer (0.15 M, pH 5.0) after which H_2O_2 (MilliporeSigma, USA) was added to a final concentration of 0.0015%. A total of 150 µl of this TMB solution was then added to each well and the plate was stored at room temperature in the dark for 30 mins. The presence of anti-BACE1-GST antibody was then detected through the measurement of absorbance at both 652 nm using a SpectraMax ABS plate reader (Molecular Devices, USA).

2.3.4 *Purification of anti-GST and anti-BACE1 IgY*

A western blot was performed as before, using 1.5 mm thick gels instead of the 1 mm thick gels used prior. The BACE1, GST solution obtained from off column thrombin cleavage was used, before the GST had been removed. The maximum amount of solution was added to each well. After transferral of the protein from the PAGE gel to the PVDF membrane (Pall Corporation, USA), the membrane was activated again by wetting it with methanol. Using the molecular weight marker as a guideline, strips of the blot were cut so as to separate the GST and BACE1 bands respectively. The PVDF strips were incubated in acidic glycine buffer (100 mM glycine, pH 2.5) for five mins. This buffer was discarded and washed twice with TBS-Tween for five mins each. The membrane strips were then blocked with TBS containing 0.5% (w/v) BSA for one hour at room temperature with gentle rocking. The blocking solution was then removed and the strips washed with TBS-Tween as before. Crude antibody samples isolated from all eggs collected from week seven of immunization were diluted 1 in 5 in TBS-BSA. The membrane strips were then incubated in this antibody solution overnight at 4°C with gentle shaking. This solution was then removed and the strips washed twice again with TBS-Tween followed by a further two washes in PBS. Bound antibodies were then eluted by adding 1 ml acidic glycine buffer, followed by a ten min incubation at room temperature. This buffer, containing the eluted antibodies was then collected. The pH of this eluent was brought to a pH of 8.0 through the addition of 100 µl Tris buffer (1 M, pH 8.0). This elution process was

repeated once, bringing the final volume of each antibody solution acquired to 2.2 ml. Antibody concentrations were then determined through the measurement of A_{280} as before. To confirm the presence and specificity of these antibodies, western blots were carried out against both rBACE1 and GST.

2.4 Differential scanning fluorimetry

2.4.1 *Thermal shift assay optimization*

Expression of BACE1-GST was carried out as previously, however a further three column volumes of PBS not containing Triton X-100 were used to wash the column with cleavage being carried out in PBS. On-column thrombin cleavage was then carried out with the resulting protein being used for the DSF assays. Thermal shift assays were carried out using the high-resolution melting (HRM) mode of the Corbett RG 6 000 thermal cycler (Corbett Research, USA) using the Rotor-Gene software (Corbett Research, USA). The concentration of the fluorescent protein stain, SYPRO orange (Invitrogen, USA) was first optimized. This was carried out by using concentrations of 1X, 5X, 10X and 20X dye with a concentration of 1 µM of rBACE1 (Wright *et al*., 2017). Further, several buffers were tested including sodium acetate buffer (50 mM sodium acetate, pH 4.5), PBS as well as acetate, 2-(N-Morpholino) ethanesulfonic acid (MES), Tris (AMT) buffer (100 mM acetate, 100 mM MES, 200 mM Tris-HCl) over a range of pH values (pH 4.0, 5.0, 6.0, 7.0, 7.2, 7.4, 7.6, 7.8, 8.0). All thermal shift reactions were carried out with a reaction volume of 25 µl, with measurements being carried out between 25 and 95°C. Gain was optimized at the beginning of each experimental run. The BACE1 specific inhibitor, Verubecestat, was used as a confirmation for the thermal shift assay with the Verubecestat being used at a concentration of 1 µM. Thermodynamic parameters were determined for BACE1 from the DSF data using the following equations as per Wright *et al*., (2017). Equation 1 was used to determine the F_{max}, the predicted fluorescence intensity when the protein in question is fully unfolded. F_{Tm} corresponds to the fluorescence measured at the recorded Tm and the F_{min} corresponds to the lowest fluorescence achieved during the DSF assay. Equation 2 was used to determine the proportion of folded protein (P_f) present at each temperature with F being the fluorescence measured at the respective temperature. Equation 3 was used to calculate the proportion of unfolded protein (P_u) at each temperature. The equilibrium constant of unfolding (K_u) was then determined using equation 4 allowing for the determination of $\Delta_u G$ (equation 5). A plot of $\Delta_u G$ against temperature in Kelvin at each temperature was then made, with the line of best fit being extrapolated to 298 K to determine

the $\Delta_u G°$. Equation 6 was then used alongside the $\Delta_u G°$ value to determine the standard entropy of unfolding ($\Delta_u S°$) and subsequently the standard enthalpy of unfolding ($\Delta_u H°$).

$$F_{max} = \left(F_{T_m} - F_{min}\right) + F_{T_m} \ (1)$$

$$P_f = 1 - \frac{F - F_{min}}{F_{max} - F_{min}} \ (2)$$

$$P_u = 1 - P_f \ (3)$$

$$K_u = \frac{P_u}{P_f} \ (4)$$

$$\Delta_u G = -RT \ln K_u \ (5)$$

$$\Delta_u S = \frac{\Delta_u G}{T_m - T} \ (6)$$

$$\Delta_u H = T_m \Delta_u S \ (7)$$

2.4.2   Screening of BACE1 against ChemBridge compound library

The 50 compounds of the in-house compound library obtained from the ChemBridge DIVERSet compound collection (ChemBridge, USA) were screened against 1 µM rBACE1. The rBACE1 was mixed with 20 µM of each compound and incubated on ice for 30 mins before the addition of dye. A control using a concentration of 2% dimethyl sulfoxide (DMSO, MilliporeSigma, USA) was also carried out. Binding was detected as a shift in the T_m of the rBACE1, characterized as the temperature at which the rate of increase in fluorescence was highest.

2.5 Expression of BACE1 in a mammalian cell line

2.5.1   Generation of a BACE1 expressing cell line

HEK-293 cells were obtained through the NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH: HEK-293 Cells from Dr. Andrew Rice (catalogue number: 103). These cells were thawed and subsequently grown and maintained in growth media (90% (v/v) Dulbecco's Modified Eagle's Medium (DMEM, 10% (v/v) foetal bovine serum (FBS) and 0.2× penicillin-

streptomycin (PenStrep) all purchased from Thermo Fisher Scientific, USA) at 37°C in 5% CO_2 in T75 flasks. Cells were then washed twice with 10 ml sterile PBS and exposed to three ml 0.25% (w/v) trypsin for two mins at room temperature. After visually confirming all cells had been successfully detached from the flask surface, 3 ml growth media was added. The resulting mixture was then centrifuged at 1000 *xg* for 5 mins with the supernatant being discarded. The cell pellet was resuspended in growth media and 5 x10^5 cells were seeded into each well of a 24 well-plate and grown until 90% confluence under identical conditions as before. Once these cells reached 90% confluence, DNA solutions containing 5 µg BACE1-pcDNA-DEST 40 plasmid, 5 µl PLUS reagent (Thermo Fisher Scientific, USA) in 250 µl Opti-MEM media (Thermo Fisher Scientific, USA) were made. Volumes of lipofectamine including 2, 3, 4 and 5 µl were then diluted in 50 µl each, with each solution then being mixed 1:1 with the DNA solutions made prior and incubated at room temperature for five mins. Each resulting lipofectamine-DNA solution was then added to duplicate wells of the previously grown HEK-293 cells, at a final volume of 50 µl (final concentrations per well: 500 ng DNA, 1-2.5 µl lipofectamine). A mock-transfected control was carried out, as well as a transfection using the empty pcDNA-DEST 40 plasmid. The cells were then incubated as before for three days, after which the media was discarded and replaced with selection media (90% (v/v) DMEM, 10% (v/v) FBS, 0.2X PenStrep and 200 µg/ml G418). Growth in the presence of selection media was taken as confirmation of transfection. Once confluent, cells were transferred to T75 flasks and maintained in selection media.

2.5.2 *Assaying compound cytotoxicity in BACE1 expressing HEK-293 stable cell line*

BACE1 expressing HEK-293 cultures were grown to confluency in selection media and subsequently trypsinized with the resulting cell pellet resuspended in one ml selection media as previously. A sample of cell suspension was mixed 1:1 with Trypan blue after which cell counts and viability were determined using a Bio-Rad TC20 Automated Cell Counter (Bio-Rad Laboratories, California, USA). Each well in a 96-well plate was then seeded with 2×10^4 cells/well. Cells were incubated for four hrs at 37°C in 5% CO_2 to allow for the cells to adhere. Serial dilutions of test compound ranging in concentration from 200 µM to 1.56 µM were then added with the cells then being incubated for a further four days at 37°C in 5% CO2. After incubation, 40 µl of 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium salt (MTS) reagent containing phenazine ethosulphate (PES) was added to the 200 µl of media in each well, followed by further incubation for one hr. The absorbance at 490 nm was then recorded using a SpectraMax ABS Plus microplate reader

(Thermo Fisher Scientific, USA). The concentrations at which cytotoxicity resulted in 50% cell survival (CC_{50}) were calculated for each compound by plotting the non-linear fit of the data acquired using the log(inhibitor) vs normalized response – variable slope function in GraphPad Prism 8.4.1 (2019, GraphPad Software, USA)

2.5.3 Cellular thermal shift assay analysis

After trypsinization of confluent cultures, cells were harvested by centrifugation at 1000 *xg* for five mins. For initial western blot analysis of the transfected cells to confirm BACE1 expression, the cell pellet was washed twice with 1 ml PBS and then resuspended in 600 µl radioimmunoprecipitation assay (RIPA) buffer (150 mM NaCl, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris, 1x protease inhibitor cocktail, 1 µl/ml DNase 1). This was incubated for one hour at 4°C with gentle agitation, after which the sample was centrifuged at 10 000 *xg* for 30 mins at 4°C. The pellet was discarded and the supernatant analysed*via*western blot as per section 2.3.2. This was carried out on a transfected cell line, the mock transfected cell line as well as the cell line transfected with the empty pcDNA-DEST 40 plasmid.

In the case of the CETSA, prior to harvesting, cells were exposed to the test compounds for 2 hrs at 37°C with 5% CO2 (Almqvist *et al*., 2016). The cells were exposed to 10 µM of test compound, in the case of the positive control, 1 µM Verubecestat (Merck & Co., USA) purchased from MedChemExpress (USA) (Scott *et al*., 2016) or 0.035% DMSO in the case of the control. After incubation, cells were harvested and cell pellets washed twice with PBS. The cell pellets were then resuspended in 900 µl PBS and split into 100 µl aliquots. Seven samples were incubated at a temperature range of 40-65°C for 6 minutes using a T100 Thermal cycler (Bio-Rad Laboratories, USA). The two control samples were incubated on ice and at room temperature respectively. Samples were then allowed to cool to room temperature for 6 mins, followed by centrifugation at 5 000 xg, for 5 mins and the supernatant discarded. Each cell pellet was then resuspended in 50 µl RIPA buffer (MilliporeSigma, USA) for 15 mins at 4°C with gentle agitation. After lysis, samples were centrifuged at 10 000 xg for 20 mins at 4°C, with the resulting supernatant being analysed by SDS-PAGE, followed by western blotting as per Section 2.3.2. using anti-BACE1 rabbit IgG (MilliporeSigma, USA, SAB2108415-100UL) (1:1000) followed by goat anti-rabbit IgG conjugated to horse radish peroxidase (1:10 000)(MilliporeSigma, USA, AP156P). Visualization was carried out using enhanced chemiluminescent substrate (ECL) (Clarity ECL reagent, Bio-Rad Laboratories, USA) and

viewed with a G:BOX Chemi XR5 (Syngene, India) in the GeneSys software (2012). Densitometry was then carried out on the bands corresponding to monomeric BACE1 in all samples using the ImageJ 1x software (Schneider *et al*., 2012). The relative intensity values were then normalized into percentages with the largest intensity for each sample assigned a value of 100% and the lowest assigned as 0%. Data was plotted using a Boltzmann sigmoidal function calculating the T_{agg} using GraphPad Prism 8.4.1.

2.6 Zebrafish embryo toxicity test

All zebrafish studies were carried out with ethics approval (AREC/029/019) being granted for the project entitled: "Establishment of Zebrafish models for the study of human disease and toxicology" on 3 June 2020.12.03 attached in Appendix A. Zebrafish (*Danio rerio*) were sustained in tanks of aged water at a temperature of 28°C under the illumination of blue light (wavelength ~450 nm). Fish were fed three times per day with TetraPro Energy fish flakes (Tetra, Germany) at regular six hourly intervals. Male fish were mixed with females at a ratio of 1:2 the afternoon prior to breeding with collection reservoirs being placed in each tank. Breeding was allowed to take place for 30 mins after the onset of light. Fertilized embryos were then taken, washed using E3 medium (5 mM NaCl, 0.17 mM KCl, 0.33 mM $CaCl_2$, 0.33 mM $MgSO_4$, 5% Methylene Blue) and subsequently separated into 24-well plates at 10 embryos/well. A positive control using 4 mg/l 3,4-dichloroaniline (MilliporeSigma, USA) was used alongside negative controls containing DMSO as well as a control containing E3 media only. Embryos were exposed to compounds over a concentration range at three hrs post-fertilization (hpf) and incubated at 28.5°C. Concentration ranges were determined using a range-finding test whereby embryos were exposed to compounds at concentrations of 100 mg/L, which were subsequently diluted until 100% survival was recorded. Images of embryo development for each embryo were taken at 24, 48, 72, 96 hours post fertilization and five days post fertilization (dpf). Lethality and proportion of embryos presenting abnormalities were recorded and used to determine LC_{50} values as well as teratogenicity for test compounds. For teratogenicity, the concentration at which no abnormalities in embryo development was determined using an established scoring system based on the appearance of nine morphological endpoints. These endpoints include somite formation, notochord formation, tail structure, fin structure, heart shape, brain segmentation, facial structure and jaw structure and pharyngeal arch structure. A score up to 5 was assigned based on the level of abnormality from the correct morphology with 5 being allocated to correct morphology and a minimum score of 0.5 indicating a total lack of the structure in question, resulting in a maximum possible score of 45

in accordance with Panzica-Kelly *et al.* (2010). Whether a compound was teratogenic or not was determined by calculating the ratio between the LC_{50} and the EC_{50} both of which were calculated using a nonlinear regression using GraphPad Prism 8.4.1 (2019, GraphPad Software, USA). If this teratogenic index was greater than 1 the compound was deemed teratogenic whereas all compounds scoring teratogenic indexes less than 1 were denoted as not being teratogenic (Panzica-Kelly *et al.*, 2010) (Alafiatayo *et al.*, 2019).

2.7 Computational analysis of the BACE1 protein structure

The Schrodinger Biologics suite (2018-2) (Schrodinger, USA) was utilized for all *in silico* analysis. Docking experiments were carried out using the 3D human BACE1 protein structure 5HU1. To prepare the protein structure, the protein preparation wizard was used. All water molecules found further than 5 Å from the bound ligand were deleted, heteroatom states were generated using Epik with a chosen pH of 7.0. Missing hydrogens were added, bond orders assigned and disulphide bonds generated. After this, hydrogen bond assignments were optimized and the resulting structure was then minimized. Ligands to be screened were drawn using the 2D sketcher tool and were prepared using the LigPrep function to add missing hydrogens and generate possible conformations of the ligands. At most 32 conformations were generated per ligand and Epik was used to generate protonation states. The prepared BACE1 protein was then used to create a receptor grid through the use of the receptor grid generation tool. No constraints were used. The resulting receptor grid model was validated through redocking of its cognate ligand back into place. The root-mean-square-deviation (RMSD) value of the conformation with the highest Glide score against the original, bound ligand was calculated with the model being validated by an RMSD value of one or less. Screening of the in-house compound library against the prepared receptor grid was initially carried out using standard precision (SP), after which all hits scoring a Glide score below -5 were then screened again, using extra precision (XP) (Kumar *et al.*, 2017). All resulting hits with a Glide score less than -5 were assessed individually to determine which amino acids were involved in binding, with the ligands that demonstrated bonding to either of the catalytic aspartate residues being selected as successful hits. This was carried out on the BACE2 crystal structure 2EWY as well. Further binding sites on the BACE1 protein were determined using the SiteMap tool, with residues found in the discovered binding site. A receptor grid was generated for any potential binding site as before, however without the redocking steps, with screening taking place only

on XP setting, with all conformations exhibiting a Glide score less than -5 being accepted as binders.

2.8 In Silico screening of the Enamine REAL database for BACE1 binding compounds

The structure of compounds detected to bind BACE1 in the DSF screening, CETSA and molecular modelling simulations were then used to screen the Enamine REAL database (Enamine, Ukraine) for similar compounds using the MadFast Similarity Search (ChemAxon, Hungary, https://www.enaminestore.com/search accessed last on 14/06/2019). Compounds possessing similarity scores of greater than 0.6 in the database were then screened *in silico* for BACE1 binding as per Section 2.8. The compound scoring the lowest Glide score was then used to search for more similar compounds with this process being carried out a total of five times. The compound detected to have the highest propensity for binding BACE1 as indicated by the lowest Glide score at the end of the fifth screening was then assayed for intracellular binding of BACE1 using the CETSA as per Section 2.6.3.

Test compounds were also assessed for predicted drug likeness alongside other relevant traits using the in silico tools SwissADME, Aggregate advisor and OSIRIS property explorer. Aggregate advisor was accessed*via*http://advisor.bkslab.org/ (Irwin *et al.*, 2015) in December 2020. SwissADME was accessed*via*http://www.swissadme.ch/ last accessed in December 2020. OSIRIS property explorer was accessed*via*https://www.organic-chemistry.org/prog/peo/ last accessed in December 2020.

2.9 Statistical analysis

All statistical analyses were carried out using PSPP (GNU Project, USA). Significance is represented with a p-value < 0.05 using a two-tailed Student's T-test.

CHAPTER 3: RESULTS

3.1. Confirmation of presence of pGEX4T-1 containing the rBACE1 gene in transformed *Escherichia coli* BL21 DE3 cells

Prior to plasmid synthesis, the human BACE1 gene was analysed for possible restriction endonuclease sites using NEBcutter (Appendix B). Based on this analysis it was decided that the gene would be synthesized between the cleavage sites of EcoR1 and BamH1. The protein sequence was analysed using PeptideCutter to determine possible cleavage sites of common proteases (Appendix C). Neither Factor Xa nor thrombin cleavage sites were present and thus the BACE1 gene was synthesized in a pGEX-4T1 vector as this would allow for the removal of the GST tag using thrombin. Competent *E. coli* BL21 DE3 cells were successfully transformed with the BACE1$_{43-454}$–GST plasmid. This was demonstrated by the capability of the transformed *E. coli* to grow on 2xYT agar plates in the presence of 100 µg/ml ampicillin. To ensure the *E. coli* contained the plasmid with the full BACE1$_{43-454}$ gene insert, the plasmid was isolated from the transformed bacteria using a GeneJet plasmid DNA isolation kit. The plasmid DNA was isolated at a concentration of 296.3 ng/µl with an A$_{260}$/A$_{280}$ ratio of 1.99, indicating pure DNA to be present. This DNA was exposed to a restriction digestion by the restriction enzymes EcoR1 and BamH1 with both the digested and undigested DNA being visualized as bands on a 1% (w/v) agarose gel shown in Figure. 3.1. Four distinct bands resulted from the restriction digestion of the DNA with EcoR1 and BamH1, present at 1 200, 1 800, 2 400 and 5 000 base pairs (bp) respectively as seen in Figure 3.1. This indicates the release of the rBACE1 gene insert (1 200 bp) from the vector (5 000 bp; lane 1). These bands indicate that contaminating plasmid DNA may have been present within the transformed bacteria. This was further substantiated by the presence of more than three bands in the undigested DNA sample. The undigested DNA appears in seven bands ranging from 3 500 to greater than 10 000 bp in size (lane 2). Due to this anomalous plasmid DNA the isolated plasmid was also sent for gene sequencing (Inqaba Biotec, RSA) which confirmed the presence of the BACE1 gene (GeneID: 23621) as found on GenBank (Appendix D).
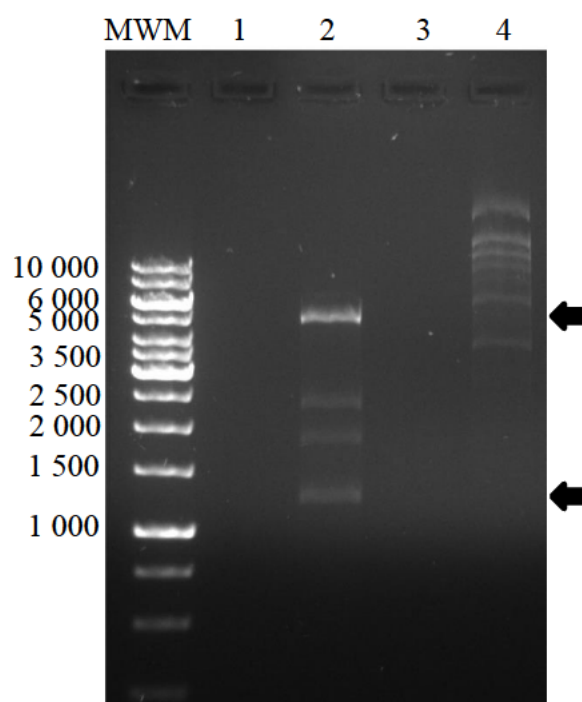
Figure 3.1: Plasmid DNA isolation from *E. coli* BL21 DE3 cells transformed with a BACE1-pGEX4t-1 plasmid evaluated on a 1% (w/v) agarose gel. The OGene 1kb Ruler molecular weight marker was run in the lane designated MWM. Samples were run as follows: lane 1: no sample, lane 2: DNA isolated from *E. coli* BL21 DE3 cells after restriction digestion using EcoR1 and BamH1. Lane 3: no sample, lane 4: undigested DNA isolated from *E. coli* BL21 DE3 cells. Arrows indicate bands corresponding to the 5 000 base pair linear pGEX4T-1 and the 1 200 base pair rBACE1 gene. The gel was stained using ethidium bromide and viewed under ultraviolet light.

3.2. Recombinant expression of the rBACE1-GST fusion protein

3.2.1. *Optimization of rBACE1-GST expression*

The transformed *E. coli* overexpressed a protein with a size of approximately 70 kDa after induction with IPTG, however this protein appears to be produced, to a lesser degree, in the uninduced sample as well, indicating the possibility of leaky expression. In all samples, this protein was more prominent in the insoluble fractions. Both the soluble and insoluble fractions are compared in the SDS-PAGE gels seen in Figure. 3.2.
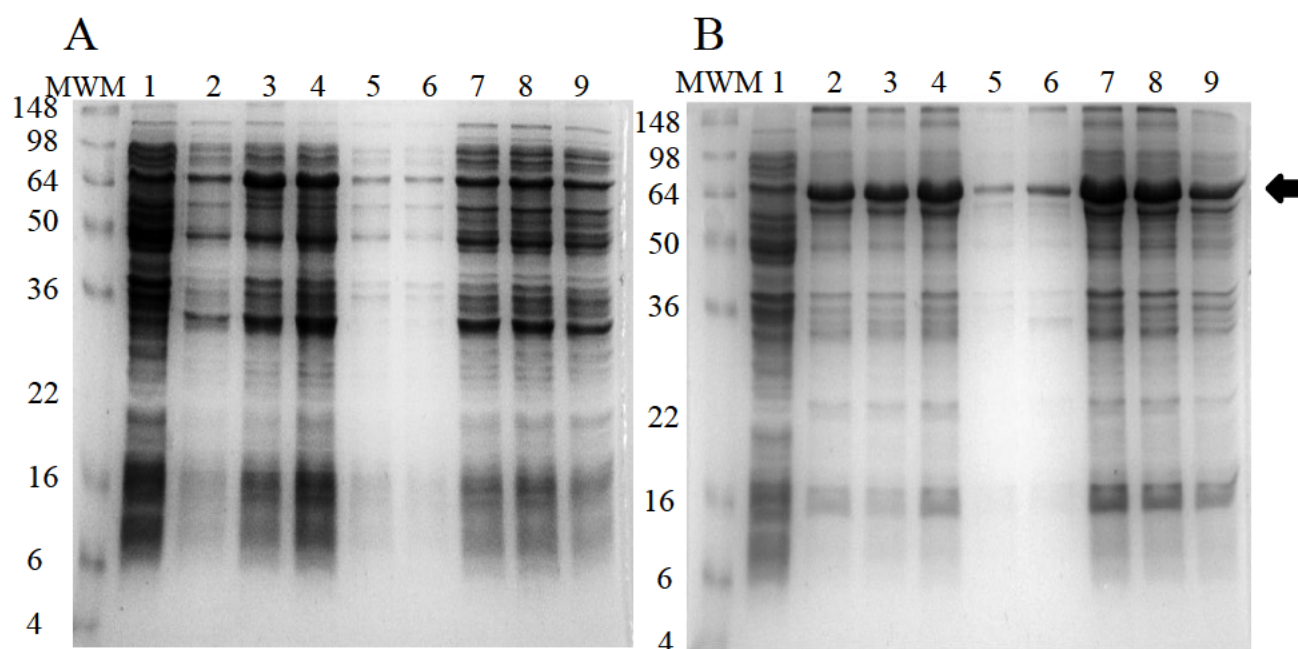
Figure 3.2: Optimization of the expression and solubility of the BACE1-GST fusion protein in *E. coli* BL21 DE3. Protein samples were separated and visualized through non-reducing SDS-PAGE using 12.5% gels, stained with Coomassie blue R250. The total amount of protein expressed by *E. coli* BL21 DE3 transformed with the BACE1-pGEX4T-1 plasmid under varying expression conditions were separated into the soluble protein fraction (A) and insoluble protein fraction (B). All cultures were induced overnight at 16°C, in 2xYT using 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) unless stated otherwise. The SeeBlue Plus2 Pre-stained protein standard was run in the lane designated MWM. Samples were run as follows: lane 1: cell pellet from uninduced *E. coli* BL21 DE3 culture, lane 2: samples from a 16°C induction overnight in 2xYT media, lane 3: sample from culture induced in LB media, lane 4: sample from culture induced with 0.5 mM IPTG, lane 5: sample from culture induced for 4 h at 16°C, lane 6: sample from culture induced at 37°C for 4 h, lane 7: sample from culture induced in the presence of 1% ethanol, lane 8: sample from culture induced using 0.5 mM IPTG in the presence of 1% ethanol, lane 9: sample from culture induced in the presence of 3% ethanol. The arrow indicates the overexpressed protein at ~70 kDa.

Less overall protein was generated from a four-hour induction at both 37 and 16°C (67% and 58% less respectively) (lanes 5 and 6 of Figure 3.2 respectively), as compared to induction overnight at 16°C (lane 1 of Figure 3.2.), this did not however result in increased solubility of the overexpressed protein. The LB media, as well as the sample induced with 0.5 mM IPTG, run in lanes 3 and 4 respectively, resulted in the highest level of expression of the soluble form of this protein from the *E. coli*. Even in these samples however, the majority of protein remains insoluble. The largest bands found in the insoluble fractions are from the induction with 0.5 mM IPTG as well as induction in the presence of 1% ethanol, although combining the two

conditions resulted in a decrease in the intensity of this band. The presence of 3% ethanol resulted in less expressed protein than expression in the presence of 1% ethanol. Measuring relative intensities of these bands confirmed that the insoluble fraction produced during induction in the presence of 1% ethanol did indeed result in the largest amount of protein as seen in Table 3.1. It was noted that a band in the uninduced sample was present at a similar size to the rBACE1-GST and as a result densitometry on a western blot was carried out to ensure densitometry was carried out solely on the rBACE1-GST.

Table 3.1: Relative intensities of rBACE1-GST bands present on SDS-PAGE and western blot analyses over the course of rBACE1 expression optimization.

	Relative band intensity[a]			
	Soluble fraction		Insoluble fraction	
Expression conditions	SDS-PAGE[b]	Western blot	SDS-PAGE	Western blot
16°C O/N[c]	15431.36	1654.678	30069.28	13280.81
LB[d] media	25929.67	8846.134	28541.74	12655.57
0.5 mM IPTG[e]	24273.89	6151.477	33067.57	14940.4
16°C 4hr	6349.698	1498.385	9111.012	2122.87
37°C 4hr	4917.577	1054.849	14267.08	7355.426
1% ethanol	21251.6	1541.971	39061.59	16035.13
1% ethanol 0.5mM IPTG	24158.31	2783.506	37681.57	15034.76
3% ethanol	19056.31	2585.799	24235.62	10378.79

[a] Relative intensities determined using the ImageJ 1x image analysis software (Schneider *et al.*, 2012).

[b] Sodium dodecyl sulphate-polyacrylamide gel electrophoresis

[c] Overnight

[d] Luria broth

[e] Isopropyl β-D-1-thiogalactopyranoside

The overexpressed protein was confirmed as the rBACE1-GST fusion protein through the use of western blot (Figure. 3.3). The chicken anti-GST IgY utilised detected multiple bands in all samples. The likelihood of non-specific binding being the cause of these bands was ruled out by the absence of these bands in the uninduced sample. These anomalous bands may indicated degradation of the rBACE1-GST protein over the course of expression. Induction in LB media as well as induction in the presence of 0.5 mM IPTG resulted in the most soluble rBACE1-GST

represented by the prominent bands at 70 kDa in lanes 3 and 4 of Figure 3.3, panel A. In the remaining samples, minimal soluble protein was detected. The insoluble fractions however, displayed large amounts of protein present at 70 kDa. Densitometry of the western blot confirmed the findings of the SDS-PAGE gels indicating that, despite the band present in the uninduced samples, the results determined from the densitometry of the SDS-PAGE gels were reliable. It was also confirmed the insoluble fraction of the sample obtained from induction in the presence of 1% ethanol indeed contained the most rBACE1-GST. Thus, induction at 16°C overnight in the presence of 1 mM IPTG was selected for all future experiments. This was chosen as, although both LB media and 0.5 mM IPTG resulted in more soluble protein exhibiting relative band intensities of 25929.67 and 24273.89 respectively, the largest amount of rBACE1-GST was present in the 1% ethanol insoluble fraction allowing for the greatest possible protein yields with a relative band intensity of 39061.59. Due to the insolubility of the rBACE1-GST, an attempt to reduce insolubility of the protein by disrupting aggregation was attempted. The insoluble fractions obtained from expression were exposed to various conditions known to inhibit protein aggregation. After separation of the insoluble and soluble fractions, the insoluble fraction was resuspended and sonicated to disrupt remaining protein aggregates. An SDS-PAGE gel was then run on the resulting samples (Figure. 3.4). None of these conditions resulted in increased solubility however, indicating that the insolubility observed would need to be resolved during a refolding step.
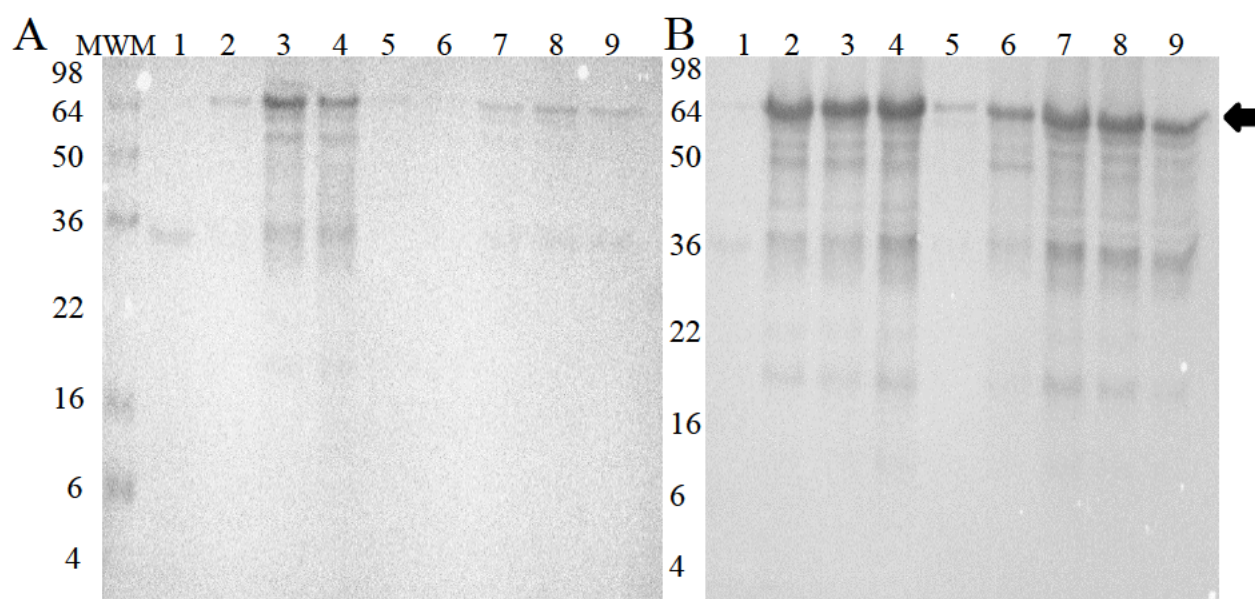
Figure 3.3: Western blot analysis of the soluble (A) and insoluble (B) proteins produced during optimization of the BACE1-GST fusion protein expression in *E. coli* BL21 DE3. Samples were electrophoresed using non-reducing SDS-PAGE in 12.5% gels and transferred to polyvinylidene fluoride membranes. The membrane was probed with chicken anti-GST IgY primary antibodies (supplied by professor THT Coetzer) and goat anti-chicken IgY secondary antibodies conjugated to horse radish peroxidase (Advansta, USA). Results were visualized using enhanced chemiluminescence substrate. Samples were separated into the soluble protein fraction (A) and insoluble protein fraction (B). All cultures were induced overnight at 16°C, in 2xYT using 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) unless stated otherwise. The SeeBlue Plus2 Pre-stained protein standard was run in the lane designated MWM. Samples were run as follows: lane 1: cell pellet from uninduced *E. coli* BL21 DE3 culture, lane 2: samples from a 16°C induction overnight in 2xYT media, lane 3: sample from culture induced in LB media, lane 4: sample from culture induced with 0.5 mM IPTG, lane 5: sample from culture induced for 4 h at 16°C, lane 6: sample from culture induced at 37°C for 4 h, lane 7: sample from culture induced in the presence of 1% ethanol, lane 8: sample from culture induced using 0.5 mM IPTG in the presence of 1% ethanol, lane 9: sample from culture induced in the presence of 3% ethanol. The arrow indicates the overexpressed protein confirmed to be rBACE1-GST.
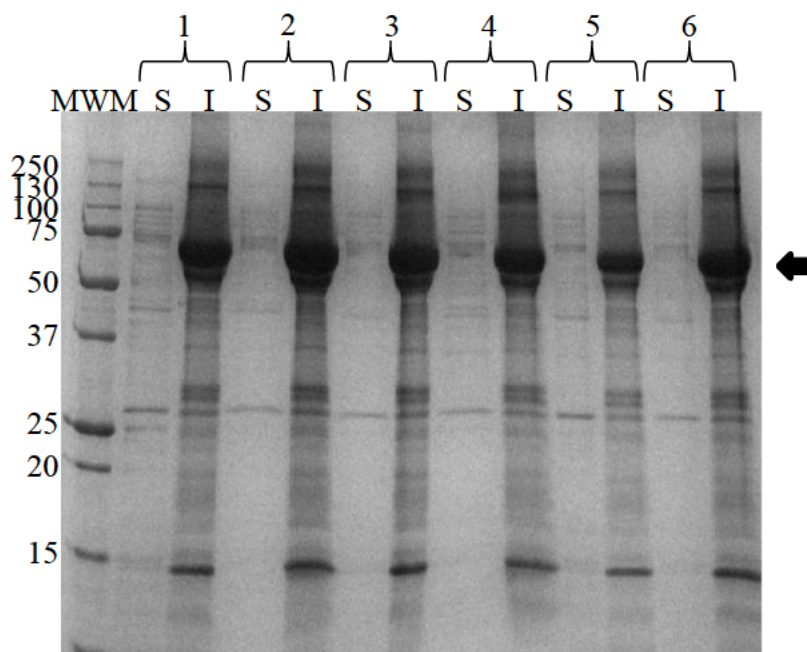
Figure 3.4: Use of various conditions to inhibit aggregation of the overexpressed rBACE1-GST fusion protein. Samples from the insoluble fraction of *E. coli* BL21 DE3 cells, were exposed to conditions known to reduce aggregation, sonicated and separated into insoluble and soluble fractions through centrifugation. These fractions were run on an SDS-PAGE gel and stained with Coomassie blue R250. The Precision Plus Protein molecular weight marker was run in the lane designated MWM. Soluble fractions (S) were run alongside the corresponding insoluble fractions (I). The samples run were as follows: 1: control sample with no added aggregation inhibitors, being sonicated and separated into soluble and insoluble fractions, 2: sample was exposed to 1% polyethylene glycol 6 000, 3: sample was exposed to 10% glycerol, 4: sample was exposed to Triton X-100 at a final concentration of 1%, 5: sample was exposed to 200 mM L-arginine, 6: sample was exposed to 200 mM sucrose. Arrow indicates the rBACE1 protein band at ~70 kDa.

3.2.2. Purification of rBACE1-GST

Expressed rBACE1-GST protein was solubilized and refolded using 10% sarkosyl solution. Samples were collected at each step of the purification and analysed on a gel and western blot shown in Figure. 3.5. Degradation of rBACE1-GST was evident throughout the course of purification. After affinity purification however, only three protein bands were observed in lanes 6, 7, 8 and 9 of Figure 3.5, panel A. The band corresponding to the rBACE1-GST at approximately 70 kDa, a band at 65 kDa and a third band at 15 kDa. The rabbit anti-BACE1 IgG did not detect the latter protein, it did however detect the two larger proteins. This indicates that this smallest protein could be a section of degraded GST-tag. A band is also present in lanes 2, 3, 4 and 5 in panel B of Figure 3.5 at 25 kDa. This band is not clearly visible in panel A, nor is it present after affinity purification in lanes 6 to 9 in panel B. Refolding was also

attempted using a mixture of reduced and oxidized glutathione. This refolding step however resulted in rBACE1-GST not binding to the affinity column after dialysis. A Bradford assay was carried out on the protein samples obtained following affinity purification. A standard curve (Appendix E) was generated using stock solutions of BSA of known concentrations. Using this standard curve, it was determined that the yield of rBACE1-GST was approximately 12 mg/L of *E. coli* culture.
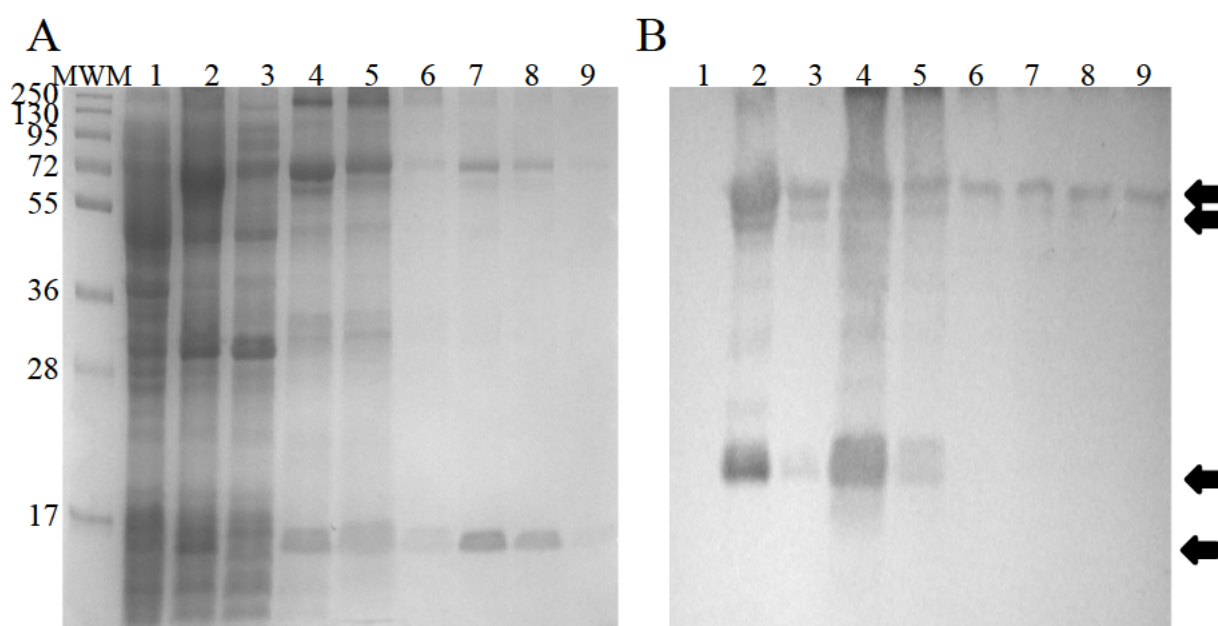
Figure 3.5: Analysis of the expression and purification of rBACE1-GST from *E. coli*. Analysis carried out using 12.5% non-reducing SDS-PAGE (A) and a corresponding western blot (B) probed with rabbit anti-BACE1 IgG primary antibody (MilliporeSigma, USA) followed by goat anti-rabbit IgG secondary antibody conjugated to horse radish peroxidase (MilliporeSigma, USA). Visualization was carried out using enhanced chemiluminescent substrate. The Pageruler Plus prestained molecular weight marker was run in the lane designated MWM. Samples were run as follows: lane 1: uninduced sample, lane 2: induced cell pellet, lane 3: soluble fraction, lane 4: insoluble fraction, lane 5: unbound fraction, lanes 6 through 9: fractions eluted from the glutathione-agarose column. Arrows indicate the protein bands at ~70 kDa, ~65 kDa, ~25 kDa and ~15 kDa.

3.2.3. Gelatine gel zymography

After isolation of uncleaved rBACE1-GST, thrombin was used to remove the GST-tag of the fusion protein. The resulting cleaved rBACE1 was then demonstrated to possess protease activity through the use of the gelatine gel zymogram in Figure. 3.6. This activity was displayed by the presence of a clear band at 50 kDa, which is similar to that of the predicted size of the cleaved rBACE1 (46 kDa) in lanes 3 of Figure 3.6. The uncleaved rBACE1-GST in lane 2 did not demonstrate any activity however, indicating that the presence of a GST-tag prevents

cleavage of protein substrate. The use of gelatine gel zymography to determine proteolytic activity of BACE1 has not been previously described in the literature. As a result, recombinant BACE1 produced in HEK-293 cells purchased from MilliporeSigma (USA, catalogue No. S4195), was used as a positive control to confirm that BACE1 is indeed capable of gelatinase activity. This BACE1, run in lane 1, exhibited protease activity at a size of approximately 110 kDa. There was however no activity at the predicted monomeric size of 70 kDa. Further activity of the mammalian BACE1 is detected at >250 kDa which may be the result of the formation of a BACE1 tetramer. Neither the activity at 110 kDa nor the activity at >250 kDa are present in the rBACE1 samples.
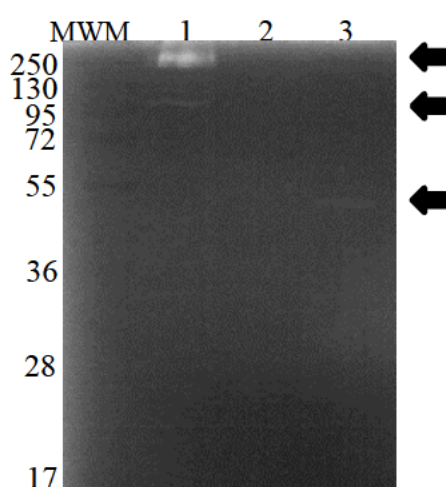
Figure 3.6: Assaying for proteolytic activity of purified rBACE1 in a gelatine gel zymogram. Electrophoresis was carried out under non-reducing conditions using a 12.5% acrylamide gel containing 0.1% gelatine. The gelatine zymogram was incubated at 37°C overnight in 50 mM sodium acetate buffer. The gel was stained in Coomassie blue R250. The Pageruler Plus prestained molecular weight marker was run in the lane designated MWM. Samples were run as follows: lane 1: commercial BACE1 produced in HEK-293 cells (MilliporeSigma, USA, catalogue No. S4195), lane 2: purified uncleaved rBACE1-GST fusion protein, lane 3: purified rBACE1. Arrows indicate detection of proteolytic activity at ~50 kDa in the rBACE1 sample and at ~110 kDa and >250 kDa in the commercial BACE1 sample.

3.3. Generation of IgY antibodies from the eggs of chickens

3.3.1. Analysis of anti-rBACE1-GST antibody production

The uncleaved rBACE1-GST protein generated previously was chosen for use in the immunization of chickens as this allowed for the production of both anti-BACE1 and anti-GST antibodies, which could then be separated during the process of antibody purification. Through

the use of an ELISA, the week in which the most antibodies against rBACE1-GST were produced was determined to be week seven (Figure 3.7). High background readings were present in the pre-immune sample. The eggs produced one week after immunization exhibited a consistent increase in TMB signal when analysed*via*ELSIA indicating an increase in the production of anti-rBACE1-GST antibodies. Signal was significantly higher than that of the pre-immune sample from week 3 onwards, up until week 11, however after week seven, the signal began decreasing, corresponding with the final booster on week six. The signal plateaued for the remaining weeks. Accordingly, anti-BACE1 and anti-GST antibodies were isolated from the eggs produced during the seventh week of immunization so as to generate the highest concentration of antibody possible.

3.3.2. Purification of anti-BACE1 and anti-GST IgY

The concentration of the antibodies affinity-purified using rBACE1 and GST bound to nitrocellulose from week seven was then determined through measurement of A_{280}. The concentration of the anti-BACE1 IgY antibodies was 0.272 mg/ml and the anti-GST IgY antibodies to be 0.254 mg/ml using an extinction coefficient of 1.25 (Goldring and Coetzer, 2003). Antibodies purified this way were then used in a western blot against a purification profile of rBACE1 to confirm specificity as shown in Figure. 3.8. One band was detected at the predicted size of the cleaved rBACE1 in the final purified sample and was also detected in the western blot. The rBACE1 ran at a size of approximately 50 kDa which is in alignment with the predicted size of 46 kDa in both the SDS-PAGE gel as well as the western blot. In panel C, lane 7, the chicken anti-GST IgY antibody did not detect this protein, further indicating successful removal of the GST-tag. In the same lane, the chicken anti-GST IgY antibody detected a protein at 20 kDa, assumed to be remaining GST-tag. This was not visible in the SDS-PAGE gel however, indicating this protein to be present at a low concentration.
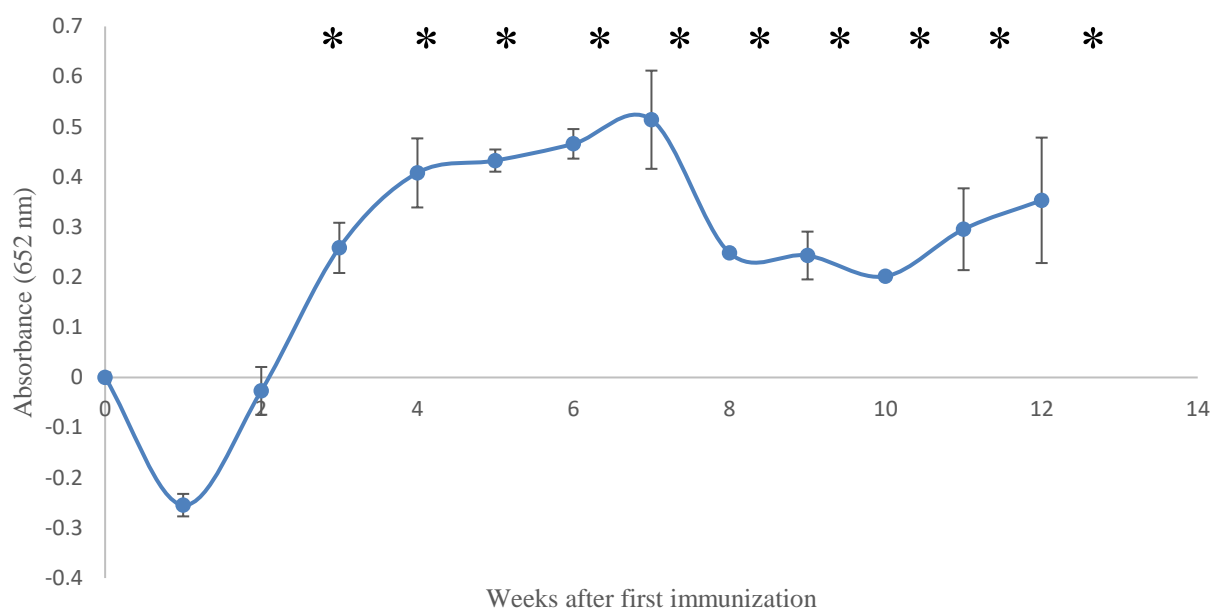
Figure 3.7: Enzyme-linked immunosorbent assay analysis of anti-BACE1-GST antibody production in chickens over the course of immunization. Microtitre plates were coated with 1 µg/ml of pure BACE1-GST. The total IgY fraction isolated from an egg from each week of immunization was used as the primary antibody at a concentration of 1 mg/ml. A goat anti-chicken IgG antibody conjugated to horseradish peroxidase was used as a secondary antibody (Advansta, USA). The substrate used was 3,3',5,5'-Tetramethylbenzidine with absorbance readings being read at wavelengths of 652 nm. Error bars represent standard deviations of triplicate samples. Samples with readings significantly higher than the pre-immune sample signal are represented *via* asterisks, with significance being determined utilizing a p-value of 0.05.

3.4. Confirmation of mammalian BACE1 detection in HEK-293 cells by purified anti-BACE1 IgY

The chicken anti-BACE1 IgY antibodies detected BACE1 in the HEK-293 cells demonstrating the ability to detect BACE1 despite post-translational modifications, as shown in Figure 3.9. BACE1 was present in the lysates of the HEK-293 cells that were transfected with the BACE1-pcDNA-DEST 40 plasmid as well as the HEK-293 cells transfected with the empty pcDNA-DEST 40 vector. The mock transfected HEK-293 cells also displayed BACE1 expression. This was expected as BACE1 is known to be expressed in HEK-293 cells (Schaefer *et al*., 2016). Protein bands at 65 and 110 kDa were present in all lysates corresponding to the sizes of the BACE1 monomer and dimer respectively. Two protein bands corresponding to sizes around 35-30 kDa were also detected by the chicken anti-BACE1 IgY antibodies. These sizes do not correspond to any known forms of BACE1 indicating that these proteins represent cleavage

products of BACE1, which is possible as the combination of their sizes result in a similar size to BACE1. Cleavage products of BACE1 at these sizes have not been reported in HEK-293 cells however. An alternative theory is that the anti-BACE1 antibody produced, binds non-specifically to these proteins, potentially with both bands representing two forms of the same detected protein.
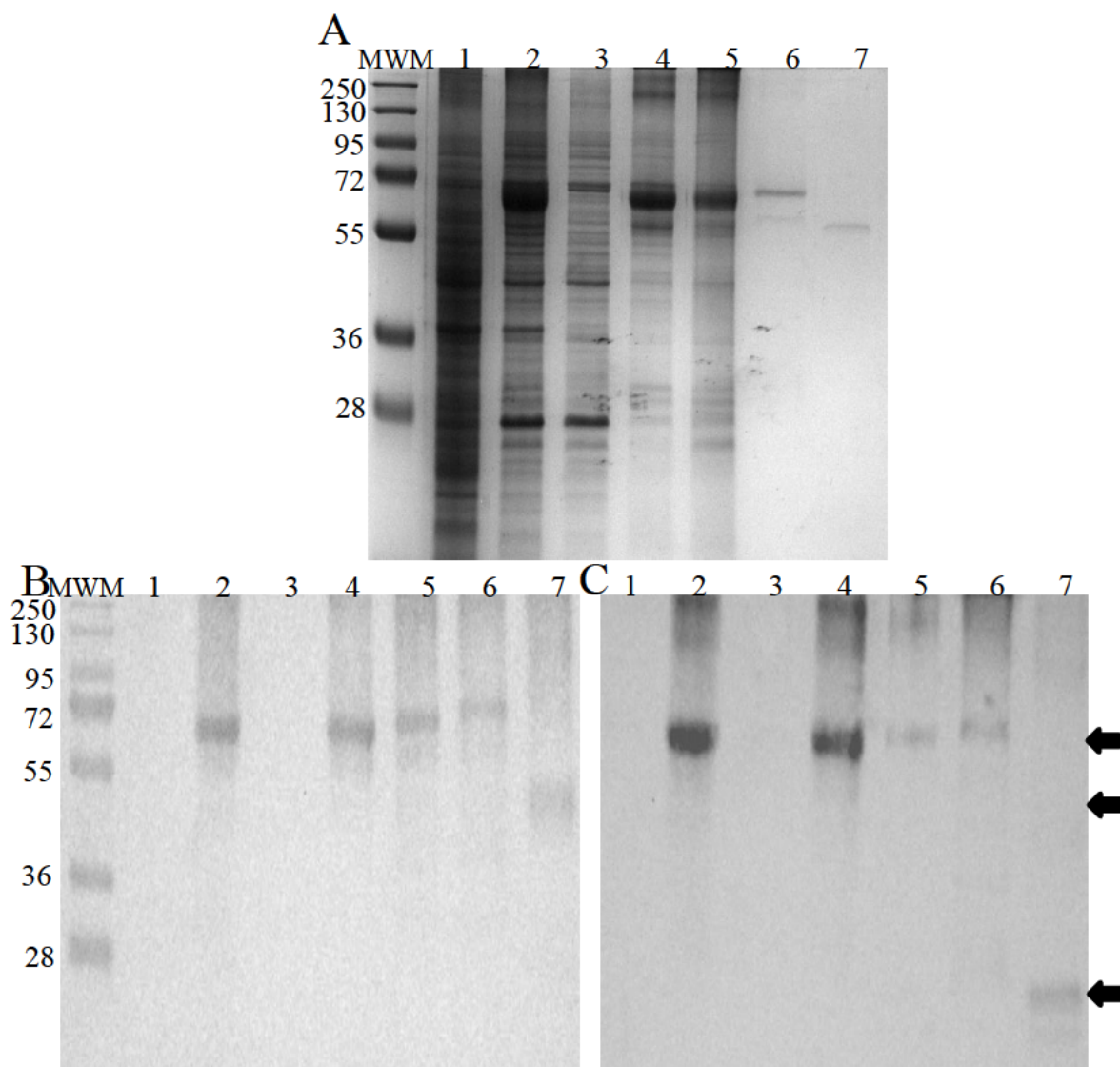
Figure 3.8: Analysis of the expression and purification of rBACE1 from *E. coli*. Analysis carried out using (A) a 12.5% non-reducing SDS-PAGE, (B) a corresponding western blot probed with chicken anti-BACE1 IgY antibody followed by goat anti-chicken IgG secondary antibody conjugated to horse radish peroxidase (Advansta, USA) and (C) a western blot using chicken anti-GST IgY antibody. Protein bands were visualized using enhanced chemiluminescent substrate. The Pageruler Plus prestained molecular weight marker was run in the lane designated MWM. Samples were run as follows: lane 1: uninduced sample, lane 2: induced cell pellet, lane 3: soluble fraction, lane 4: insoluble fraction, lane 5: unbound fraction, lane 6: uncleaved rBACE1-GST, lane 7: cleaved rBACE1. Arrows indicate the protein bands at ~70 kDa, ~50 kDa and ~20 kDa.
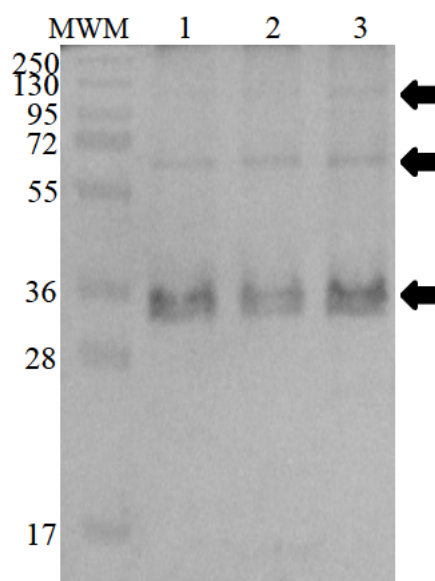
Figure 3.9: Analysis of the expression of BACE1 in HEK-293 cells transfected with the pcDNA-DEST 40 vector. Samples were electrophoresed using non-reducing SDS-PAGE in 12.5% gels and transferred to polyvinylidene fluoride membranes. The membrane was probed with chicken anti-GST IgY primary antibodies and goat anti-chicken IgY secondary antibodies conjugated to horse radish peroxidase (Advansta, USA). Results were visualized using enhanced chemiluminescence substrate. The PageRuler Plus prestained molecular weight marker was run in the lane designated MWM. Samples were run as follows: lane 1: HEK-293 cell line transfected with the pcDNA-DEST 40 vector with no BACE1 insert, lane 2: mock transfected cell line, lane 3: HEK-293 cells transfected with the pcDNA-DEST 40 vector. Arrow indicates the BACE1 protein bands at ~110 kDa, ~65 kDa and ~35 kDa.

3.5. Screening of rBACE1 against an in-house compound library using differential scanning fluorimetry

3.5.1. Determining the melting temperature of rBACE1 in the presence and absence of the specific inhibitor, Verubecestat

A DSF assay was carried out on rBACE1 in order to determine its melting temperature. This was carried out in several buffers including 50 mM sodium acetate buffer (pH 4.5), PBS (pH 7.4) as well as AMT buffer (pH 4.0, 5.0, 6.0, 7.0, 7.2, 7.4, 7.6, 7.8, 8.0). The only buffer in which the rBACE1 protein exhibited a defined peak was PBS, presented in Figure 3.10 and as a result, all DSF experiments from this point on were carried out in PBS. It was noted that previously described thermostability experiments of BACE1 at pH 7.4 have been carried out,

where it was determined that BACE1 is in an inactive conformation at this pH which resulted in increased thermostability which is potentially why a defined peak was detected at this pH. It is however unclear why no peak was detected from the rBACE1 at this same pH in the AMT buffer, although it is possible that the buffer transfer process might have resulted in denaturation resulting in a lack of a defined peak. An alternative theory is that the components of the AMT buffer may be affecting the protein structure directly as it has been seen that buffer components can affect protein dynamics in solution (Long and Yang, 2009), however this has not been reported in molecular dynamics of protein folding. The Tm value of rBACE1 was $82 \pm 0.21°C$ in the absence of Verubecestat. This is similar to the recorded Tm of 84°C of BACE1 produced in mammalian cells at pH 7.4 (Hayley *et al*., 2009). It is important to note that the Tm reported by Hayley *et al*. was an estimate of 84°C as a result of the limitations of the circular dichroism spectroscopy used, whereas the Tm detected in this study was within the limits of detection of the DSF experiment. The presence of a shift in Tm from 82 to $84.15 \pm 0.28°C$ when in the presence of Verubecestat confirmed that the rBACE1 model could be used to screen for binding molecules. This shift was confirmed to be statistically significant (p-value < 0.05), with experiments being run in duplicate. The melt curves were then used to determine the $\Delta_u G$ of the rBACE1 protein in PBS, at each temperature interval using equations 1 to 5, plotted in Figure 3.11. Data was extrapolated to 298 K to determine the $\Delta_u G°$. It was observed that the coefficient of determination (R^2) for the line of best fit was 0.934 ± 0.007 for BACE1 in the absence of Verubecestat. In the presence of Verubecestat the R^2 value was calculated at 0.733 ± 0.06 indicating that BACE1 does not undergo single-step equilibrium unfolding, while in the presence of Verubecestat. As a result, the $\Delta_u G°$ for BACE1 in the presence of Verubecestat could not be calculated. The $\Delta_u G°$ of BACE1 in the absence of Verubecestat was then used to calculate the corresponding standard $\Delta_u S°$ and $\Delta_u H°$ of unfolding, as per Wright *et al*, 2017. For BACE1 in PBS, the $\Delta_u G°$ was calculated to be 112 ± 0.1 kJ/mol, the $\Delta_u S°$ was 1.82 ± 0.002 kJ/mol K and $\Delta_u H°$ was 642 ± 0.7 kJ/mol. These values are higher than those of the proteins studied by Wright *et al*, such as the values reported for carbonic anhydrase, chymotrypsin, lysozyme and peroxidase. The values of which ranged from 20.7 ± 0.3 to 60.4 ± 0.6 kJ/mol for $\Delta_u G°$, $0.63 \pm 0.0.01$ to 1.60 ± 0.02 kJ/mol K for $\Delta_u S°$ and 212 ± 5 to 536 ± 6 kJ/mol for $\Delta_u H°$ values, all of which are lower than the values reported in this study for rBACE1, indicating high thermostability in PBS.
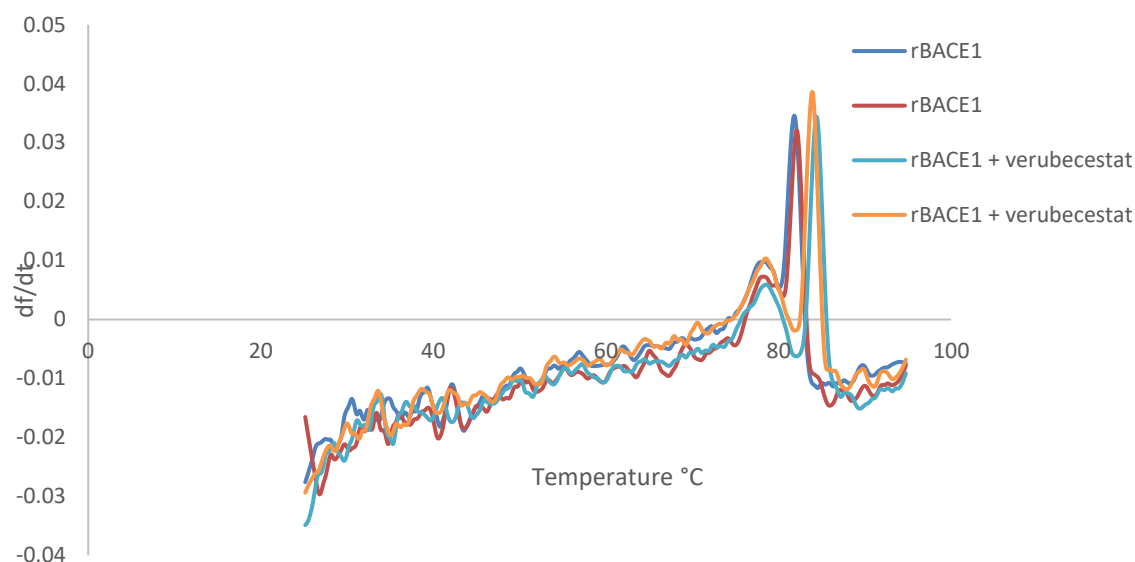
Figure 3.10: The melting profile of 1 μM rBACE1 in the presence and absence of 1 μM of the BACE1 specific inhibitor, Verubecestat. The plot of the derivative of the relative fluorescence of SYPRO orange dye in the presence of the rBACE1 protein only, as well as rBACE1 in the presence of Verubecestat over the temperature range of 25-95°C is displayed. Samples were assayed in duplicate.
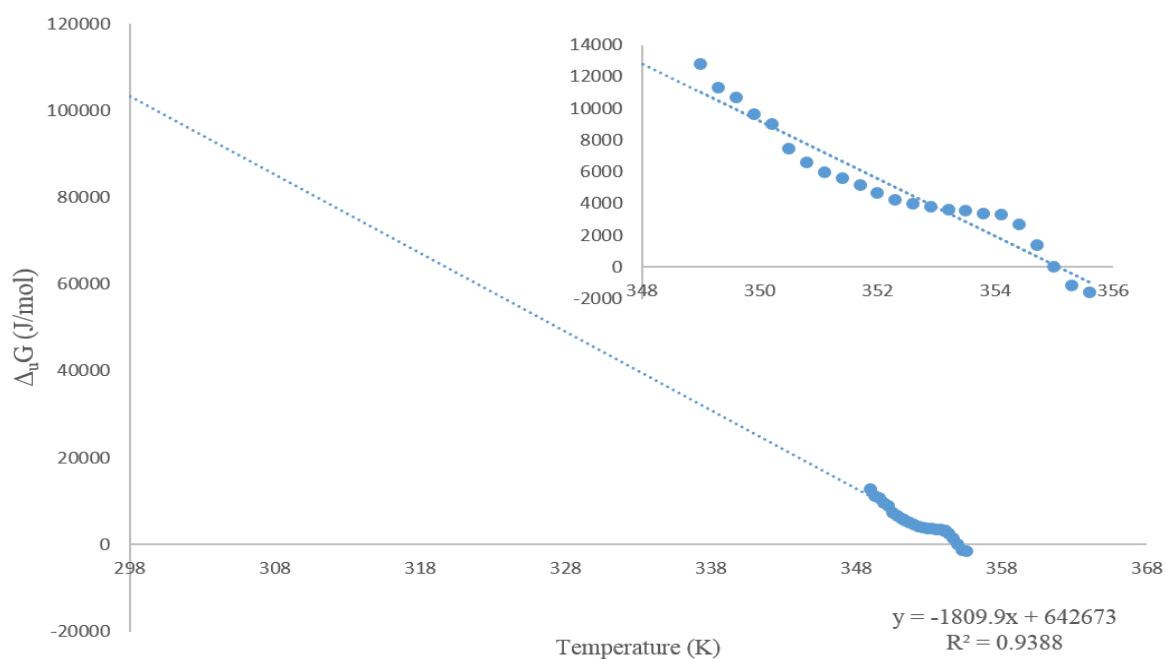
Figure 3.11: The change in Gibbs energy of unfolding of BACE1 during protein unfolding over the course of a differential scanning fluorimetry assay. Gibbs energy ($\Delta_u G$) was calculated using the fluorescence values from the linear phase of the melt curve. Values were then extrapolated to 298 K to determine the standard Gibbs energy of unfolding ($\Delta_u G°$). The equation for the line of best fit is: y=-1809.9x + 642673. The coefficient of determination (R2) was equal to 0.9388. The insert highlights the region of the graph generated by the differential scanning fluorimetry data.

3.5.2. Screening of the ChemBridge compound library subset against rBACE1

The rBACE1 was screened against 50 compounds from the ChemBridge compound library with the compounds at a final concentration of 20 µM during the assay, with the resulting Tm values represented in Figure 3.12. In the presence of 2% DMSO, the Tm of the rBACE1 decreased from 82°C to 75.4 ± 0.64°C indicating that DMSO has a destabilizing effect on rBACE1. This diminishing effect of DMSO on thermostability has been reported previously, where it was found to reduce the thermostability of lysozyme (Magsumov *et al*., 2019). It has been suggested by Yu *et al*, in 2016, that this effect of inorganic solvents on protein thermostability is not the result of a disruption of secondary protein structures, but the skewing of the proportion of unfolding protein states in equilibrium which results in a decreased amount of energy being required to push this equilibrium in favour of unfolded protein states (Yu *et al*., 2016). This reduction in thermostability may potentially reduce the accuracy of the assay if this equilibrium is sufficiently skewed to unfolded protein states, as it would result in an elongated, lower intensity, signal peak leading to higher deviations in compounds only capable of binding certain conformations of the partially folded protein. In contrast, compounds exhibiting binding capability for many conformations would still exhibit low variance. In the presence of both DMSO and Verubecestat, the Tm of rBACE1 increased from 75.4°C to 77.76 ± 0.15°C. It was also noted that $\Delta_u G°$ could not be calculated for BACE1 in the presence of DMSO due to lack of linearity when $\Delta_u G$ was plotted against temperature.

Ten compounds induced a statistically significant negative shift in Tm using a p-value of 0.05 as depicted in Figure 3.12. A negative ΔTm is indicative of a ligand binding to a state of the protein with lower stability than its native form (Kabir *et al*., 2016) and thus any compounds resulting in apparent negative shifts were disregarded. Overall, three compounds induced positive, statistically significant thermal shifts indicating possible binding (p-value < 0.05). The names and structures of these compounds are shown in Table 3.2 alongside Verubecestat. The ΔTm of Verubecestat was 2.355 ± 0.66°C. The ΔTm of **C19**, **C34** and **C39** were 2.4± 0.65, 3.045 ± 0.66 and 2.625 ± 0.64 °C respectively.
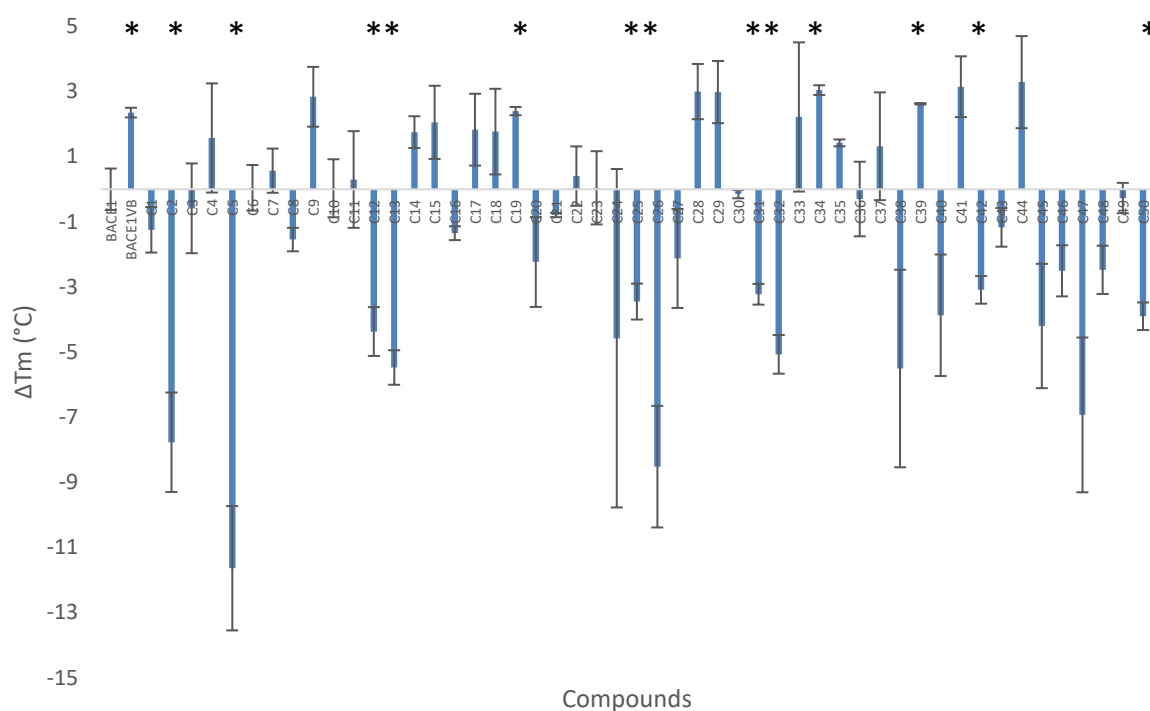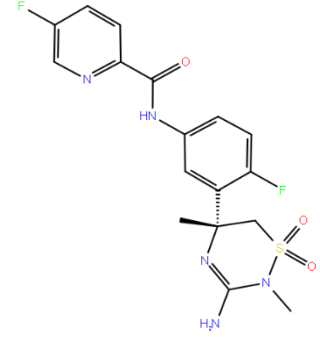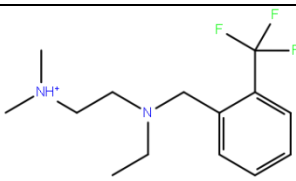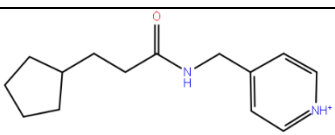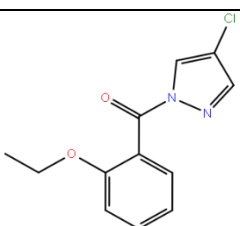
Figure 3.12: The ΔTm of rBACE1 brought about by the presence of various compounds. Test Compounds were screened against rBACE1 using differential scanning fluorimetry to determine any change in the Tm of rBACE1. Denaturation was measured through the use of the SYPRO orange dye over the temperature range between 25 and 95°C. Data is presented as the difference in melting temperature (Tm) of rBACE1 in the presence of 2 µM compound, compared to the Tm of rBACE1 of 75.4°C. Error bars represent standard deviations of duplicate samples. Samples exhibiting Tm values significantly different to the rBACE1 sample are indicated *via* asterisks, with significance being determined utilizing a p-value of 0.05.

Table 3.2: Compounds which result in a positive thermal shift in rBACE1 as detected by differential scanning fluorimetry.

Verubecestat	
N-ethyl-N',N'-dimethyl-N-[2-(trifluoromethyl)benzyl]-1,2-ethanediamine **(C19)**	
3-cyclopentyl-N-(4-pyridinylmethyl)propanamide **(C34)**	
4-chloro-1-(2-ethoxybenzoyl)-1H-pyrazole **(C39)**	

3.6 Screening for intracellular ligand binding using the cellular thermal shift assay

3.6.1 Assaying for cytotoxicity of BACE1 binding compounds in BACE1-overexpressing HEK-293 cell line

In preparation of the CETSA in the BACE1-overexpressing HEK-293 cell line, cytotoxicity of all three compounds identified in Section 3.5.2. was determined. In Figure 3.13, no compound exhibited a cytotoxic effect great enough to decrease cell viability below 50% at the concentrations used. It was noted that **C19** produced the greatest level of cytotoxicity of the

three compounds with **C34** possessing the least. This indicated that all three compounds were suitable for use in the CETSA at the test concentration of 10 μM.
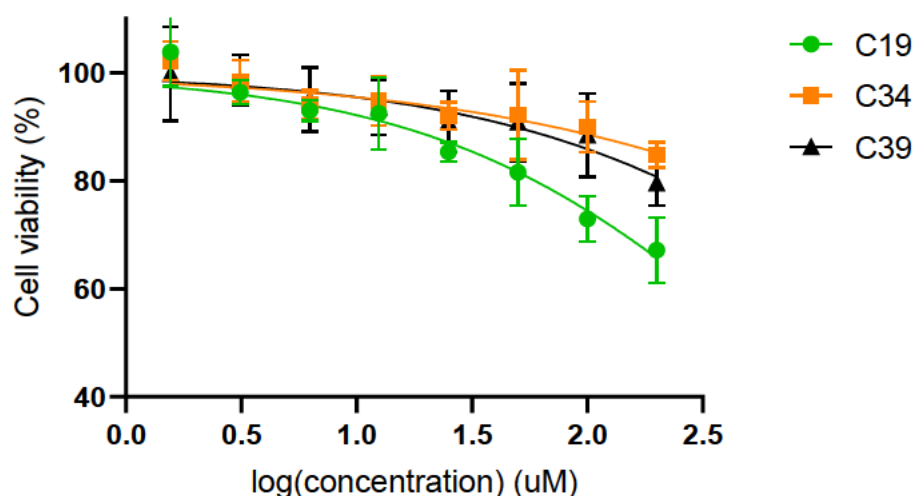
Figure 3.13: Cytotoxicity assay depicting the effect of three BACE1 binding compounds on BACE1 overexpressing HEK-293 cell viability. Cell cultures were exposed to compound concentrations ranging from 200 to 1.56 μM over the course of four days after which 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium salt (MTS) reagent containing phenazine ethosulphate was added with absorbance at 490 nm being measured to determine cell viability. Cells were exposed to MTS for three hours. Cell viability curves are indicated for compound C19 (●), C34 (■) and C39 (▲) respectively. Error bars represent standard deviations of duplicate readings.

3.6.2   Determining the T_{agg} of intracellular BACE1 produced in HEK293 cells

A CETSA was carried out on the BACE1-expressing HEK-293 stable cell line after exposure to DMSO. Proteins corresponding to sizes of >260, 130, 65 and 60 kDa were present in the soluble fractions on the western blots where all of the proteins decreased in concentration as temperature increased (Figure 3.14). With regards to the >260 kDa protein band, the temperature-dependent decrease in concentration of soluble protein confirmed that this protein band was not the result of aggregation. It was also noted that the 65 kDa protein possessed the greatest thermostability whereas the 60 kDa protein was the least thermostable, indicating that a substantial increase in thermostability is brought about by glycosylation as has been described in other transmembrane proteins (Öberg *et al.*, 2011).
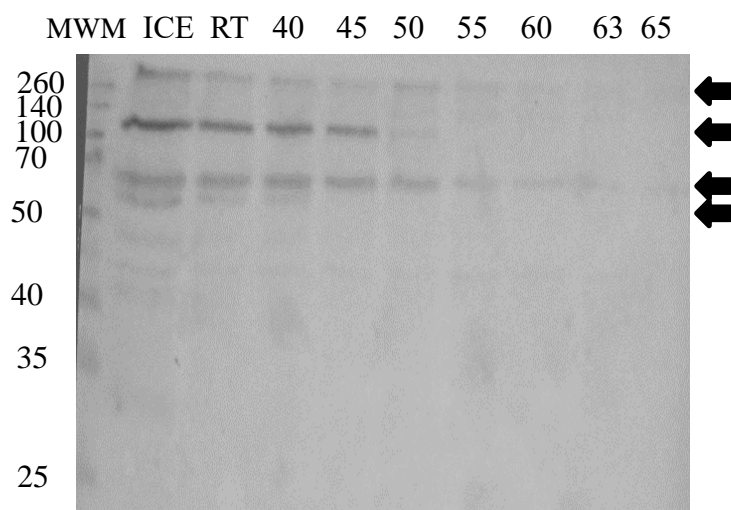
Figure 3.14: Representative western blot for the cellular thermal shift assay carried out on intracellular BACE1 overexpressed by HEK293 cells after exposure to 0.025% dimethyl sulfoxide. The western blot resulting from the cellular thermal shift assay of monomeric beta-site APP cleaving enzyme-1 (BACE1) expressed in human embryonic kidney-293 cell line (HEK293) after exposure to dimethyl sulfoxide (DMSO). The lane containing Spectra™ Multicolor Broad Range Protein Ladder (10–260 kDa) molecular weight marker (MWM), the lane containing the control sample exposed to ice (ICE) and the control sample exposed to room temperature (RT) as well as the lanes corresponding to the samples exposed to the 40 – 65°C temperature range are indicated. Arrows indicate the protein bands at >260 kDa, ~130 kDa, ~65 kDa and ~60 kDa.

The results from the CETSA analysis of BACE1 after exposure of the BACE1 expressing cell-line to the test compounds is indicated in Figure 3.15. The melt profile for the monomeric BACE1 was plotted using a Boltzmann sigmoid distribution after calculating the relative intensities of each band for each test compounds, shown in Figure. 3.15B, with the T_{agg} being calculated. The T_{agg} in the presence of DMSO alone was $49.53 \pm 0.69°C$. After it was confirmed that the CETSA was indeed capable of determining the T_{agg} of BACE1 in the HEK-293_BACE1 cell line, Verubecestat as well as the three test ligands, **C19**, **C34** and **C39**, were assayed for intracellular binding using an identical set up. In the presence of Verubecestat, the T_{agg} of monomeric BACE1 was $53.27 \pm 0.89°C$, an increase in $3.74 \pm 1.13°C$ of the T_{agg} as compared to BACE1 exposed to DMSO only. This increase was expected as it is known that Verubecestat is capable of binding BACE1 intracellularly. The resulting western blots for the test compounds displayed similar banding patterns, with regards to the monomeric BACE1, as shown in Figure 3.15A. The melt profiles were generated as seen in Figure 3.15B. The T_{agg}

values were calculated to be 49.42 ± 2.05, 53.09 ± 1.30 and 47.94 ± 3.29 for BACE1 in the presence of **C19**, **C34** and **C39** respectively. Exposure to **C34** resulted in significant increase in T_{agg} indicating engagement with BACE1 as compared to the control (p-value < 0.05). It was noted however that **C19** and **C39** yielded no significant deviation from the control.
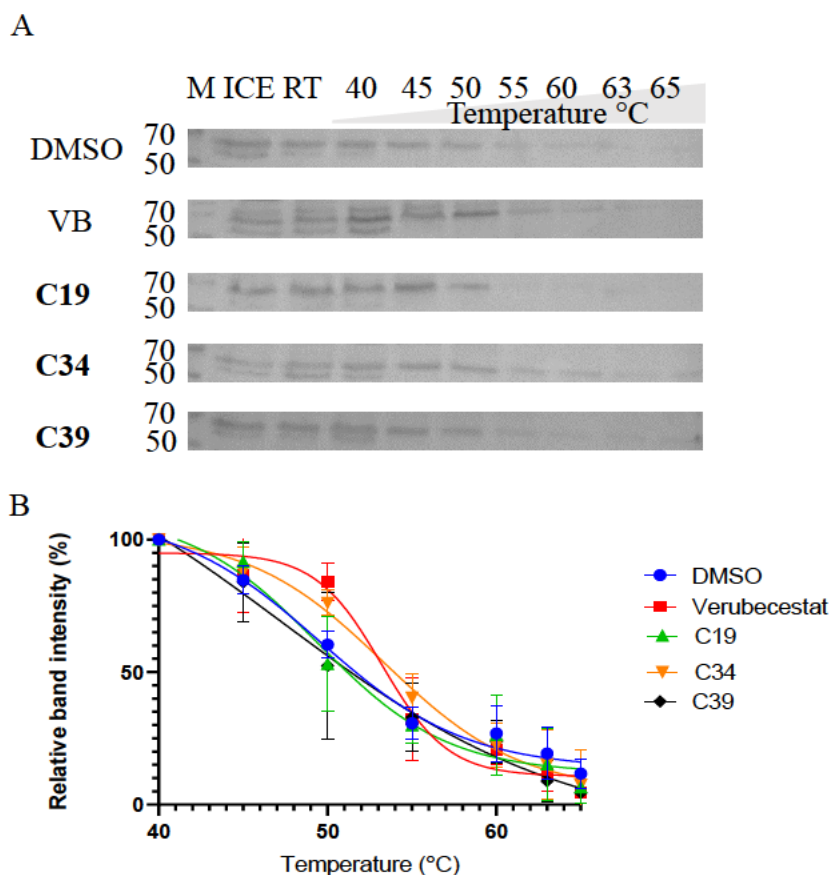
Figure 3.15: The cellular thermal shift assay of BACE1 expressing HEK293 cell line exposed to test compounds with representative western blots. Western blots f or each test compound depicting the amount of monomeric beta-site APP cleaving enzyme-1 (BACE1) produced in the human embryonic kidney-293 cell line (HEK293) over the temperature gradient of the cellular thermal shift assay (A). A representative western blot is depicted for each of the test compounds, dimethyl sulfoxide (DMSO), verubecestat (VB), N-ethyl-N',N'-dimethyl-N-[2-(trifluoromethyl)benzyl]-1,2-ethanediamine (**C19**), 3-cyclopentyl-N-(4-pyridinylmethyl)propanamide (**C34**) and 4-chloro-1-(2-ethoxybenzoyl)-1H-pyrazole (**C39**).The lane containing Spectra™ Multicolor Broad Range Protein Ladder molecular weight marker (M) control sample exposed to ice (ICE) and the control sample exposed to room temperature (RT) as well as the lanes corresponding to the samples exposed to the 40 – 65°C temperature range are indicated. The Boltzmann sigmoidal melt curves derived from plotting relative band intensity of western blot bands vs temperature (B) have also been represented for BACE1 in the presence of DMSO (●), Verubecestat (■), **C19** (▲), **C34** (▼) and **C39** (♦).

3.7. *In silico* molecular docking

3.7.1. *Generation of an in silico BACE1 model*

A BACE1 model was prepared using Schrodinger Biologics suite (2018-2) using the PDB structure 5HU1. The 3D structure of 5HU1 was chosen as it contained BACE1 co-crystallized with Verubecestat. After preparation, the model was validated through redocking, shown in Figure 3.16, with the redocked Verubecestat exhibiting an RMSD value of 0.9680 when compared with the co-crystalized Verubecestat. This was taken as confirmation that the model was sufficient for ligand screening.

3.7.2. *Screening of the ChemBridge compound library against BACE1 receptor model*

After validation, the BACE1 model was used to screen the in-house subset of the ChemBridge compound library for potential binding molecules. Initially all compounds were prepared and docked to BACE1 using the SP setting. After this, all hit compounds with Glide scores below -5 were then screened against BACE1 again using the more precise XP docking, again disregarding any hit compounds with Glide scores above -5. The interactions between these hit compounds and the amino acids of BACE1 were then assessed individually for each hit with any compounds that did not bind to either of the catalytic aspartates being disregarded. This was carried out as binding to the catalytic aspartate dyad is integral for classical inhibition of BACE1. The resulting ligands are summarized in Table 3.3 below.

Only five compounds, with a total of nine conformations, from the ChemBridge in-house library bound to either of the catalytic aspartates. Only one conformation, that of C10 (2-(3-chlorophenoxy) propanohydrazide) bound to both aspartates. Both **C19** and **C34** were detected as being potential binders to the BACE1 catalytic site, shown in Figure 3.17, corresponding to the results of the DSF experiments. It was also noted that while **C34** scored the lowest Glide score, it was detected as only binding to Asp289, whereas two conformations of **C19** were detected, each binding to a different aspartate. Compound 39 did not appear to bind at all, however, which suggests **C39** binds to BACE1 at a site other than the catalytic site. This lack of catalytic site binding calls into question the inhibitory capability of **C39**, necessitating the need for a BACE1 activity assay for further characterization, should **C39** appear to be an attractive drug candidate. When comparing C3, C10 and C23 to their corresponding DSF results, it was noted that the DSF experiment directly contrasted the docking results as C3, C10 and C23 exhibited no statistically significant change in Tm. This would indicate that C3, C10

and C23 are false positives, indicating the need for the inclusion of further steps such as molecular dynamics simulations as a confirmatory step to filter out false positives such as these.
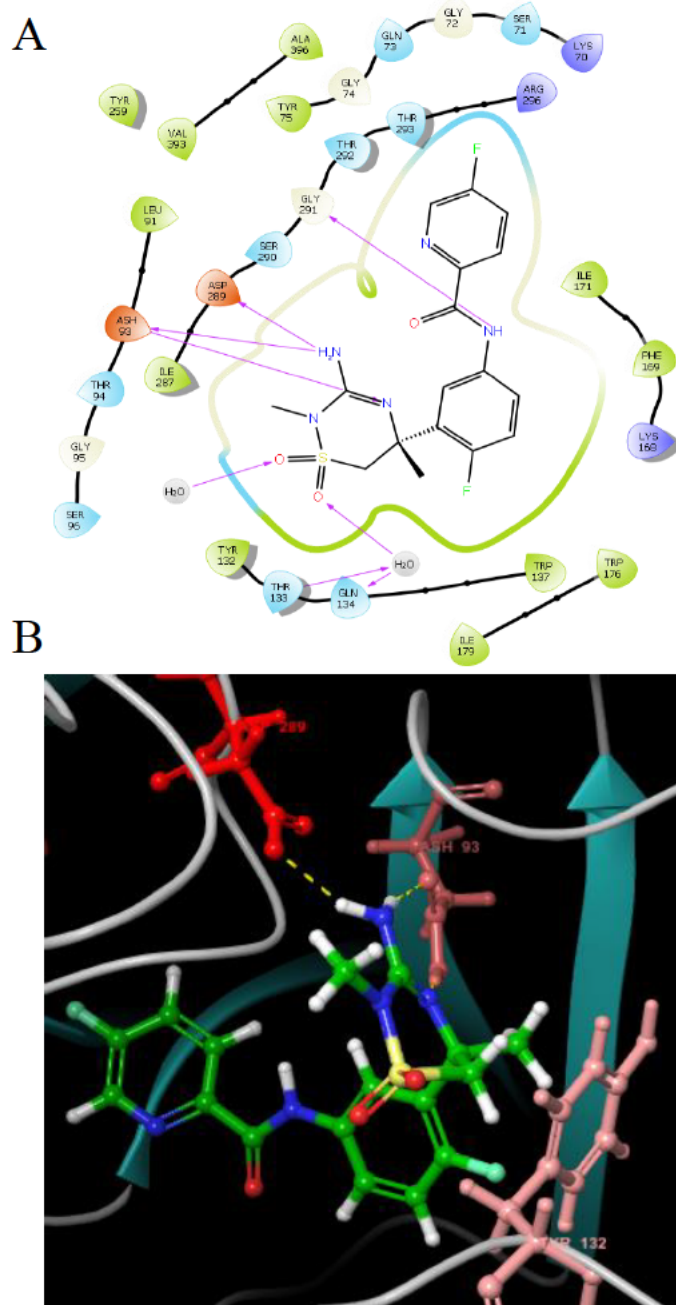
Figure 3.16: Ligand interaction diagrams of Verubecestat after redocking into BACE1 using the Schrodinger Biologics suite (2018-2). The 2D ligand interaction diagram (A) is presented alongside the 3D ligand inside the BACE1 catalytic site (B). The binding pocket is indicated by the silhouette around the ligand. Hydrogen bonds are represented by purple lines between the amino acids and the ligand. Black lines indicate the protein back bone. The ligand is coloured in green, dashed yellow lines indicate hydrogen bonds. The amino acids Aspartate 289 (top left), Aspartate 93 (top centre) and Tyrosine 132 (bottom right) are indicated in degrees of red corresponding to the level of favourable bonding occurring between these residues and the ligand, with darker shades indicating more favourable binding.

Table 3.3: Compounds predicted to be binders to the active site of BACE1 through the use of molecular docking against 5HU1 structure using Schrodinger Biologics suite (2018-2).

Compound	Glide score	Interacting amino acids
Verubecestat	-11.132	Asp93, Thr133, Gln134, Asp289, Gly291
C3	-5.566	Asp93, Gly291
C3	-5.030	Asp93
C10	-5.611	Asp93, Asp289
C10	-5.338	Asp93, Gly95
C19	-6.079	Asp93
C19	-5.993	Asp289
C23	-6.417	Asp93, Trp176
C23	-5.869	Asp93, Tyr132, Trp176
C34	-6.574	Asp289

3.7.3. Determining specificity of the BACE1 binding test compounds for the BACE1 catalytic site

As **C39** did not bind to the catalytic site of the BACE1 model, a binding site analysis was run on the BACE1 model using the SiteMap tool. This revealed a binding site on the C-terminal lobe indicated in Figure 3.18. Only **C19** was detected as a binder to this binding site, shown in Figure 3.19, presenting a Glide score of -5.490. **C34** and **C39** did not bind to this site, indicating **C39** likely neither binds to the catalytic binding site nor this binding site. Three compounds, **C19**, **C34** and **C39** were screened against BACE2 to determine if they bound selectively to BACE1 over BACE2. Only **C19** bound to the catalytic BACE2 aspartic acid dyad, with the detected binding yielding a Glide score of -6.196. This score was lower than the score resulting from the binding of **C19** to the BACE1 catalytic site indicating a potential preference of **C19** to the BACE2 enzyme. The binding of **C19** to BACE2 is indicated in Figure 3.20. The **C34** compound did not display binding, demonstrating that the in silico models of BACE1 and BACE1 predict it be specific to BACE1. **C39** displayed no affinity for either the BACE1 or the BACE2 catalytic site, however.
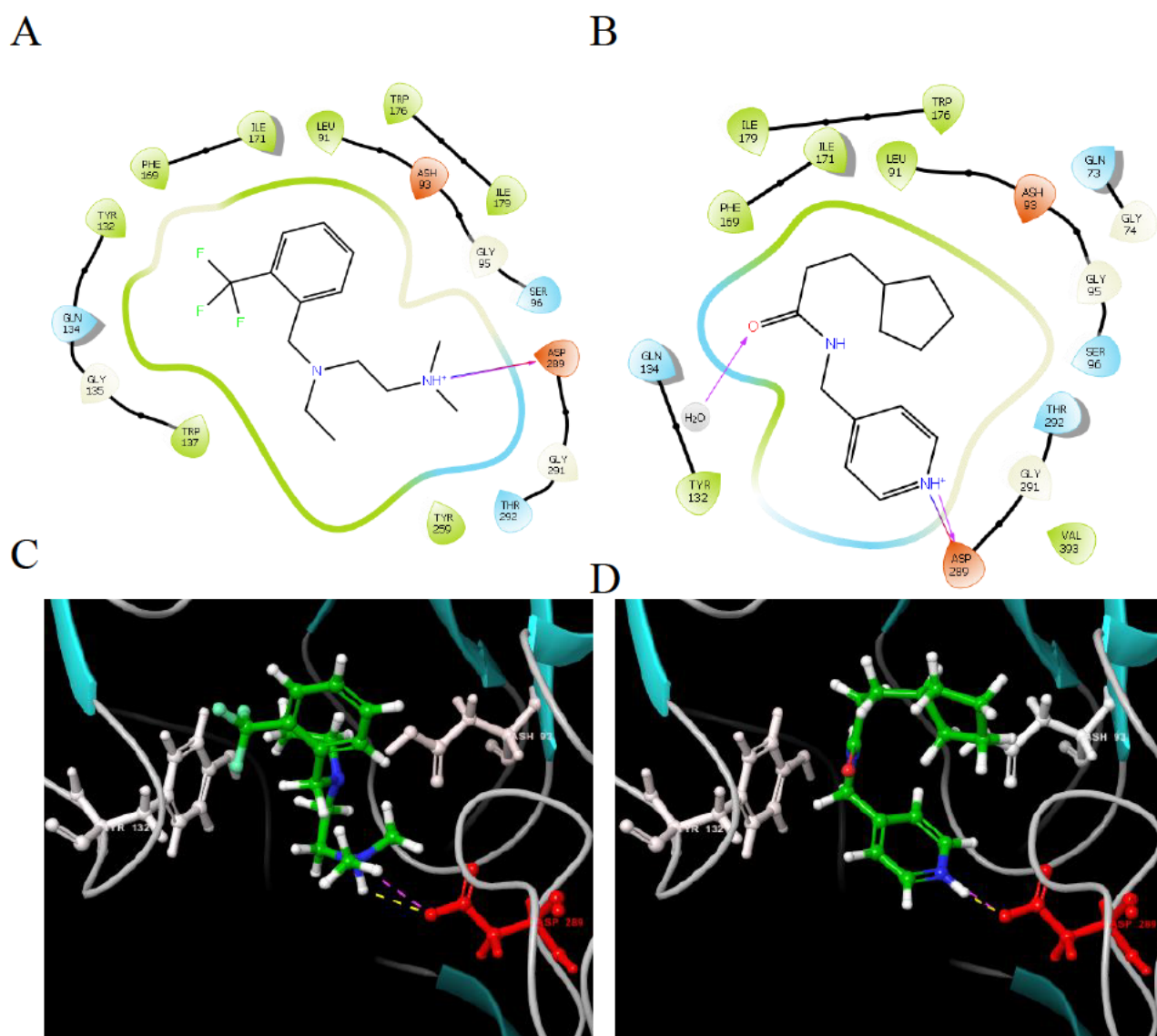
Figure 3.17: Ligand interaction diagrams of N-ethyl-N',N'-dimethyl-N-[2-(trifluoro-methyl) benzyl]-1,2-ethanediamine (C19) and 3-cyclopentyl-N-(4-pyridinylmethyl) propanamide (C34) after molecular docking against BACE1 using the Schrodinger Biologics suite (2018-2). The 2D ligand interaction diagram between N-ethyl-N',N'-dimethyl-N-[2-(trifluoro-methyl) benzyl]-1,2-ethanediamine and BACE1 is indicated in panel A. The 2D ligand interaction diagram between 3-cyclopentyl-N-(4-pyridinylmethyl) propanamide and BACE1 is indicated in panel B. The binding pocket is indicated by the silhouette around the ligand. Hydrogen bonds are indicated by purple lines between the amino acids and the ligand, blue-red lines indicate salt bridges. Black lines indicate the protein back bone. The 3D representation of N-ethyl-N',N'-dimethyl-N-[2-(trifluoro-methyl) benzyl]-1,2-ethane-diamine binding to the catalytic site of BACE1 is presented in panel C. Panel D displays the 3D representation of 3-cyclopentyl-N-(4-pyridinylmethyl) propanamide binding to the catalytic site of BACE1. The ligand is coloured in green, dashed yellow lines indicate hydrogen bonds, dashed purple lines indicate salt bridges. The amino acids Tyrosine 132 (left), Aspartate 93 (centre right) and Aspartate 289 (bottom right) are indicated in degrees of red corresponding to the level of favourable bonding occurring between these residues and the ligand, with darker shades indicating more favourable binding. Only the highest scoring conformations are demonstrated.
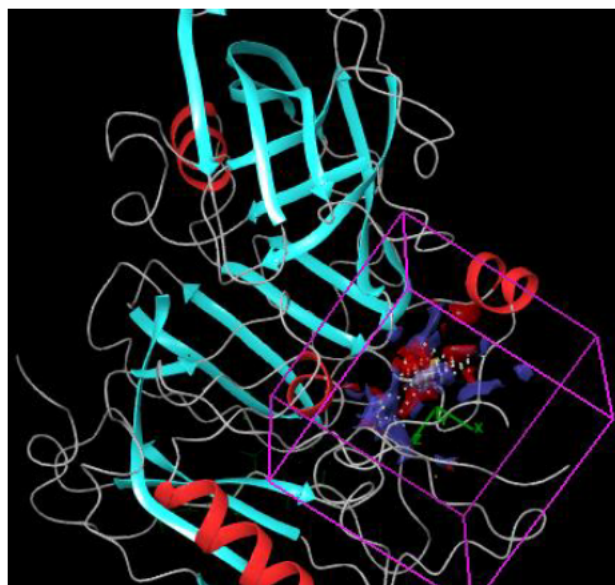
Figure 3.18: The receptor grid generated around a potential non-catalytic binding site on the BACE1 protein, found using the Schrodinger Biologics suite (2018-2). The SiteMap tool was used to detect potential binding sites on the BACE1 protein (Protein DataBank ID: 5HU1) structure. A receptor grid was generated around this site using the Receptor Grid generation tool.
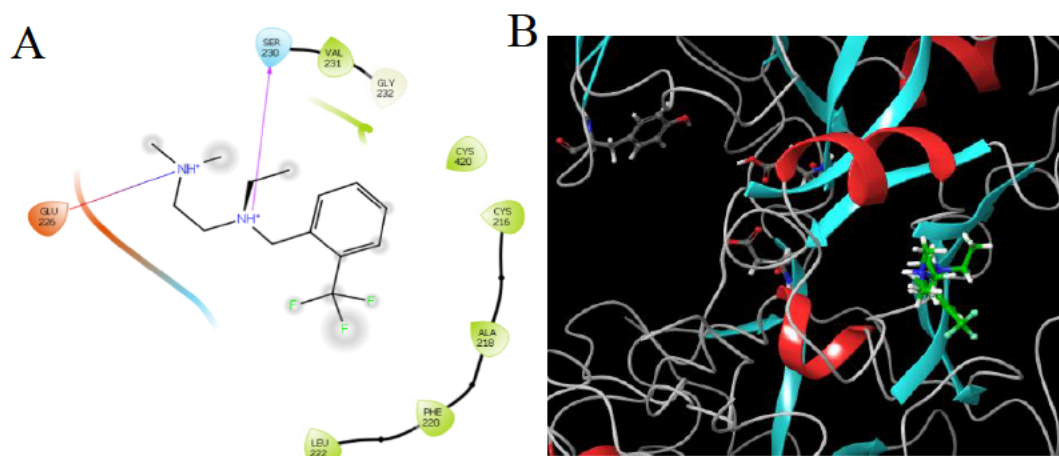
Figure 3.19: Ligand interaction diagrams of N-ethyl-N',N'-dimethyl-N-[2-(trifluoro-methyl) benzyl]-1,2-ethanediamine (C19) after docking into a potential non-catalytic BACE1 binding site, using the Schrodinger Biologics suite (2018-2). The 2D ligand interaction diagram between N-ethyl-N',N'-dimethyl-N-[2-(trifluoro-methyl) benzyl]-1,2-ethanediamine and BACE1 is indicated in panel A. The binding pocket is indicated by the silhouette around the ligand. Hydrogen bonds are indicated by purple lines between the amino acids and the ligand, blue-red lines indicate salt bridges. Black lines indicate the protein back bone. Panel B presents the 3D structure of N-ethyl-N',N'-dimethyl-N-[2-(trifluoro-methyl) benzyl]-1,2-ethane-diamine binding to the non-catalytic site of BACE1. The ligand is coloured in green. The amino acids Tyrosine 132 (left), Aspartate 93 (upper centre) and Aspartate 289 (bottom centre) are indicated.
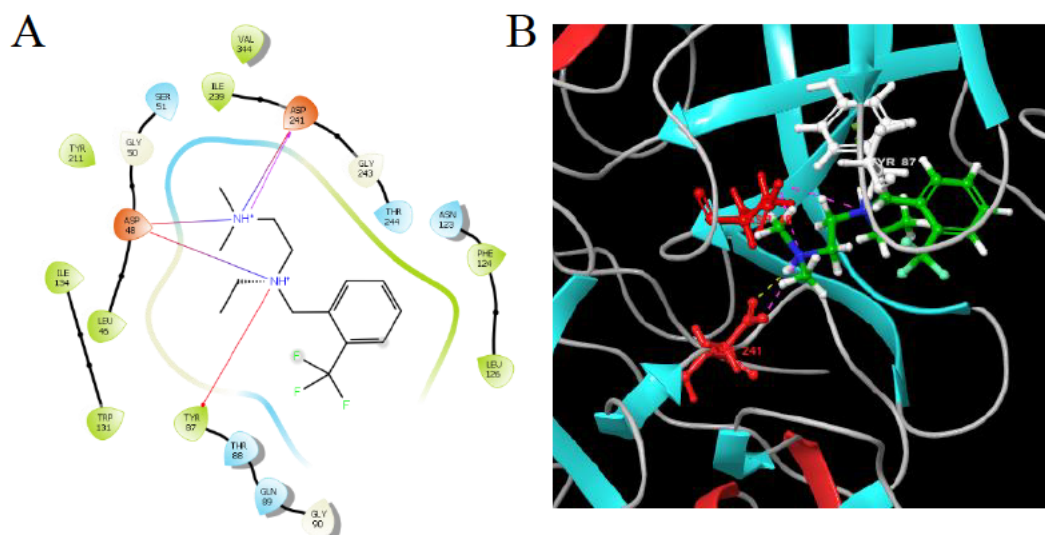
Figure 3.20: Ligand interaction diagrams of N-ethyl-N',N'-dimethyl-N-[2-(trifluoro-methyl) benzyl]-1,2-ethanediamine (C19) after docking into the BACE2 binding site, using the Schrodinger Biologics suite (2018-2). The 2D ligand interaction diagram between N-ethyl-N',N'-dimethyl-N-[2-(trifluoro-methyl) benzyl]-1,2-ethanediamine and BACE2 (Protein DataBank ID: 2EWY) is indicated in panel A. The binding pocket is indicated by the silhouette around the ligand. Hydrogen bonds are indicated by purple lines between the amino acids and the ligand, blue-red lines indicate salt bridges. Black lines indicate the protein back bone. Panel B displays the 3D representation of N-ethyl-N',N'-dimethyl-N-[2-(trifluoro-methyl) benzyl]-1,2-ethane-diamine binding to the catalytic site of BACE2. The ligand is coloured in green, dashed yellow lines indicate hydrogen bonds, dashed purple lines indicate salt bridges. The BACE2 equivalent amino acids of Tyrosine 132 (top centre), Aspartate 93 (upper left) and Aspartate 289 (bottom left) are indicated.

3.8. Screening of the Enamine REAL compound library for BACE1 binding compounds

3.8.1. Iterative in silico screening of the Enamine REAL compound database

Compound **C34** was used to screen for similar compounds capable of binding BACE1 with greater affinity than **C34**. The Enamine REAL compound database was screened for similar compounds, which were in-turn screened for BACE1 binding capability using the BACE1 model. The compound with the greatest binding affinity represented by the lowest Glide score after XP screening was then used to search for further similar molecules in an iterative manner. The results of each iteration are indicated in Figure 3.21 with the final compound after five iterations possessing a Glide score of -10.121. A CETSA assay utilizing this compound, represented by the Enamine ID Z449985064 (further referred to as **C5064**) and seen in Table 3.4, was then carried out. The number of compounds with BACE1 binding resulting in a Glide score of <-8.00 rose with each iteration of the screening process with the most significant

increase being from iteration 3 to 4 exhibiting an increase from 4 compounds to 21 compounds capable of binding BACE1 in silico. It was also noted that the percentage of the compounds screened that were capable of binding BACE1 increased from 4% to 19% from iteration 3 to 4. This is in contrast with the smaller increase from 19 to 21% seen from iteration 4 to 5.
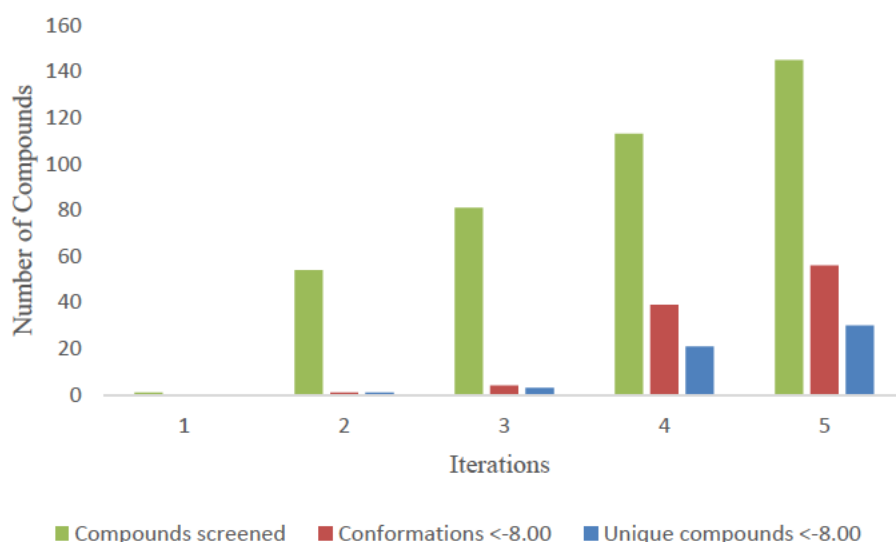
Figure 3.21: Number of hit compounds found to bind to the BACE1 catalytic site with a Glide score less than -8.00. Compounds from the Enamine REAL database were screened using the XP screening setting using the BACE1 model created in Schrodinger Biologics Suite 2018-2 (Protein DataBank ID: 5HU1). Hit compounds exhibiting the lowest Glide score were used to find similar compounds with similarity scores of 0.6 or higher using the ChemAxon addon found at https://www.enaminestore.com/search as of 06/2019. Five iterations of this process were carried out, with the number of unique compounds tested (green), the total number of conformations scoring <-8.00 per each iteration (red) and the number of unique compounds scoring a Glide score lower than -8.00 (blue) being indicated.

Table 3.4: The compound from the Enamine REAL database detected as a potential binder to the active site of BACE1, through the use of molecular docking against 5HU1 structure using Schrodinger Biologics suite (2018-2).

Glide score	Compound structure
-10.121	

3.8.2. Testing for intracellular BACE1 binding ability of enamine compound C5064

Compound **C5064** possessed several similarities to compound **C34**, most notably of which is the positively charged ring structure on one end and a non-polar ring on the other, separated by an oxygen in the centre of the compound. Cytotoxicity of **C5064** was determined, with the CC_{50} in the BACE1-overexpressing HEK-293 cell line being >200 µM. A CETSA was thus carried out, this time exposing the cell line to **C5064**. The BACE1 exposed to **C5064** during the CETSA, as seen in Figure 3.22, demonstrated a T_{agg} of 51.612 ± 1.34°C. This represented a significant increase in T_{agg} from the CETSA using DMSO of 49.53 ± 0.69°C, using a p-value of 0.05, indicating intracellular binding in the HEK-293 cell line.
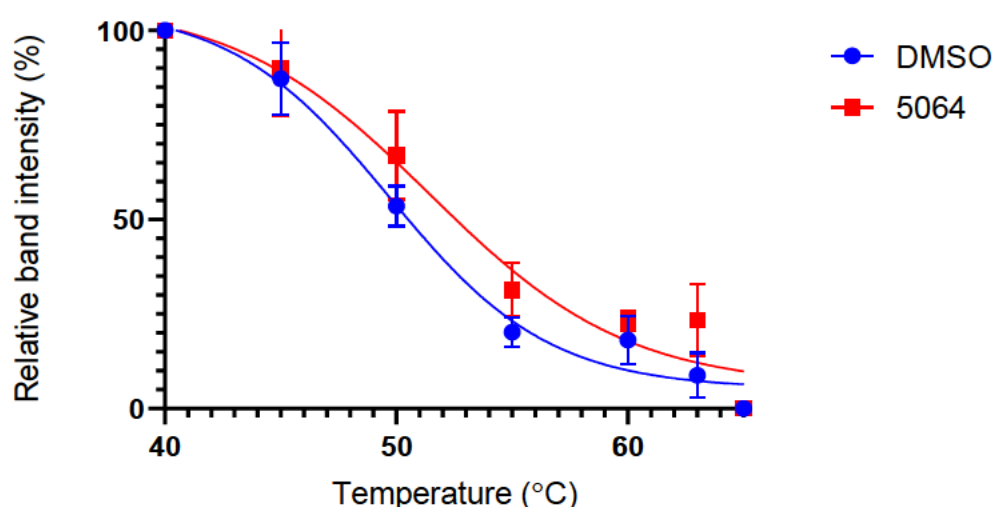
Figure 3.22: The cellular thermal shift assay of BACE1 expressing HEK293 cell line carried out after exposure of the cell line to DMSO and compounds C5064. The Boltzmann sigmoidal melt curves derived from plotting relative band intensity of western blot bands vs temperature represented for BACE1 in the presence of DMSO (●), compound **C5064** (▲).

3.9. Screening for *in vivo* toxicity using the zebrafish embryo toxicity test

3.9.1. Determining toxicity of BACE1 binding compounds using the zebrafish embryo toxicity test

In an attempt to determine *in vivo* toxicity levels of the compounds of interest in this study, a ZFET was constructed using *Danio rerio* embryos. Embryos were first exposed to test compounds at three hpf after thorough washing and resuspension in E3 buffer with survival percentage being recorded in 24-hour intervals. Exposure of the embryos to **C19** was observed to be significantly more lethal, at 7.3 µM than the control (p-value < 0.05), with 100% lethality

being observed at a concentration of 54.5 µM. It was noted that toxicity of **C19** was observable after the first 24 hours, with only exposure of 14.6 µM **C19** exhibiting a significant increase in lethality in the subsequent 96 hours. This indicates that not only is **C19** capable of penetrating the chorion, but also that it is toxic primarily to the unformed embryo, as after hatching at around 72 hours, no further increase in lethality is observed. Compound **C34** was less lethal than **C19** with significant lethality (compared to the control) observed at a concentration of 64.3 µM. A concentration of 171.4 µM, **C34** resulted in 0% survival of the embryos by day five. In contrast to **C19**, **C34** was seen to exhibit toxicity throughout the five-day incubation period with 107.1, 128.6 and 171.4 µM each exhibiting significantly lower survival percentages from 72 to 96 hpf. This increase in lethality after hatching was observed despite **C34** clearly being capable of entering the chorion, evidenced by the significant decrease in survival seen at 24 hpf, after exposure to 64.3, 85.7 and 107.1 µM. It was also noted that of the three initial test compounds that were tested, **C34** displayed the lowest toxicity with a concentration of 42.9 µM resulting in no significant decrease in survival of the embryos at day five. In contrast to **C34**, **C39** was seen to possess the greatest toxicity among the test compounds, resulting in 0% survival at a concentration of 7.85 µM, with no significant decrease in embryo survival only being detected at a concentration of 1.96 µM as compared to the control. It was also detected that this toxicity, similarly to **C19**, was only observed prior to hatching, with no change in survival being detected after 24 hours after exposure to 1.96 or 3.93 µM of **C39**. Only exposure to 7.85 µM **C39** yielded a significant decrease in survival in the 24-hour period between 24 and 48 hpf, with total lethality being observed at 72 hpf. This indicates that **C39** is highly toxic to unformed embryos but after hatching, embryos develop unimpeded.

For each compound, an LC_{50} value was calculated using a nonlinear regression function, with LC_{50} values for **C19**, **C34** and **C39** being calculated as 26.125 ± 2.45, 127 ± 9.9 and 4.52 ± 0.61 µM respectively. These values are of interest as the positive control for the zebrafish lethality test, as recommended by the OECD guidelines, is the compound 3,4-DCA, which has been reported to possess an LC_{50} value of 19.8 µM, indicating **C39** possesses a far greater toxicity to that of 3,4-DCA. **C19** was also observed to be toxic, possessing a 25% greater LC_{50} as compared to 3,4-DCA. **C34** was demonstrated to possess an LC_{50} value six-fold that of 3,4-DCA, indicating **C34** possesses far lower toxicity than that of the positive control. With regards to the Enamine compound **C5064**, no specific LC_{50} value was defined as, at the limit of solubility of **C5064** in E3 buffer (88.1 µM), no significant decrease in lethality was observed.
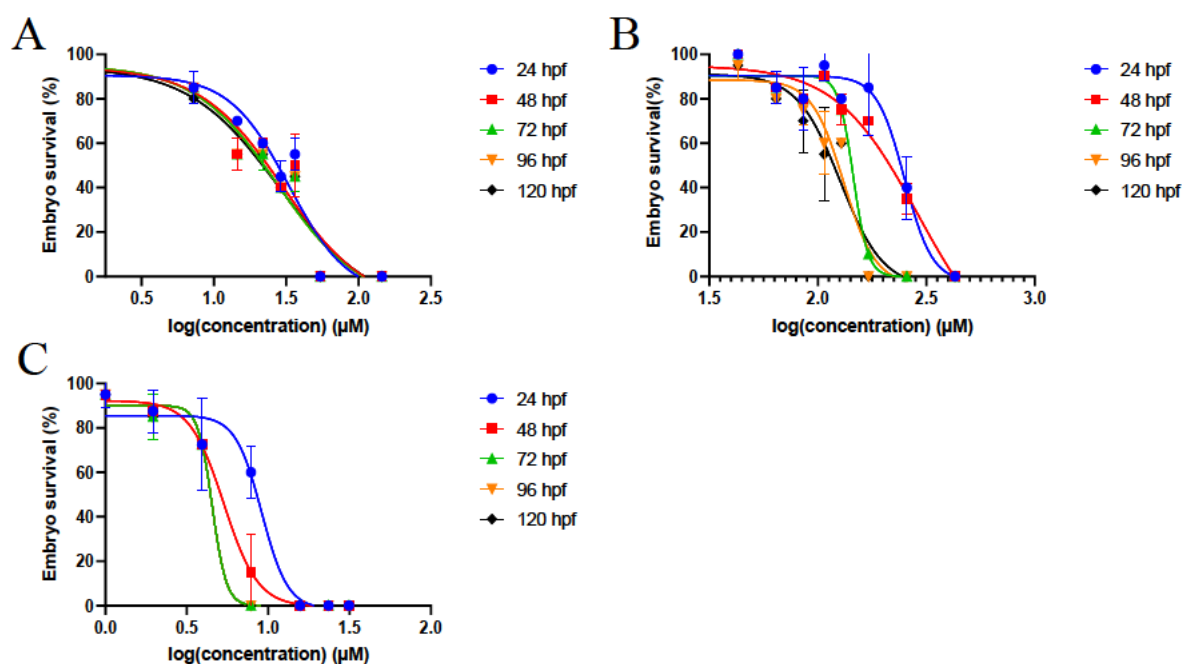
Figure 3.23: Zebrafish embryo toxicity assay depicting the toxicity of BACE1 binding test compounds at various concentrations. The survival percentage of Danio rerio embryos after exposure to various test compound concentrations at different time points are depicted. Ten zebrafish embryos per compound concentration were exposed to compounds at 3 hours post fertilization and incubated at 28.5°C for five days post fertilization (dpf) in E3 buffer. Error bars indicate the standard deviation of two biological replicate tests (n = 2). The results for compound **C19** are shown in A, with test concentrations of 7.3, 14.6, 21.9, 36.5 and 54.5 µM, **C34** displayed in B, with test concentrations of 64.3, 85.7, 107.1, 128.6, 171.4, 257.15 and 428.58 µM and **C39** in C, with test concentrations of 1.96, 3.93, 7.85, 15.70, 23.6 and 31.4 µM. Results are represented as functions of embryo survival and log (compound concentration) with graphs being represented for the time points 24 (●), 48 (■), 72 (▲), 96 (▼) and 120 (♦)hours post fertilization for each compound.

3.9.2. Establishing the zebrafish teratogenicity assay through the definition of morphological endpoints

Due to the variable nature of zebrafish strains, morphological endpoints were established for the toxicity assay in order to create references with which to detect abnormal morphologies. The structures to be assessed were the somites, notochord, tail, fins (pectoral, dorsal, caudal and ventral fins), brain, heart, facial structure (including the eye, otic capsule, facial mesenchyme and olfactory region), jaw and pharyngeal arches. Each structure scored out of a maximum of five leading to a final score out of 45. Embryos were exposed to compounds at three hpf and allowed to grow up to five dpf as seen in Figure 3.24. At five dpf morphological endpoints were captured*via*photograph and scores were assigned to each structure.
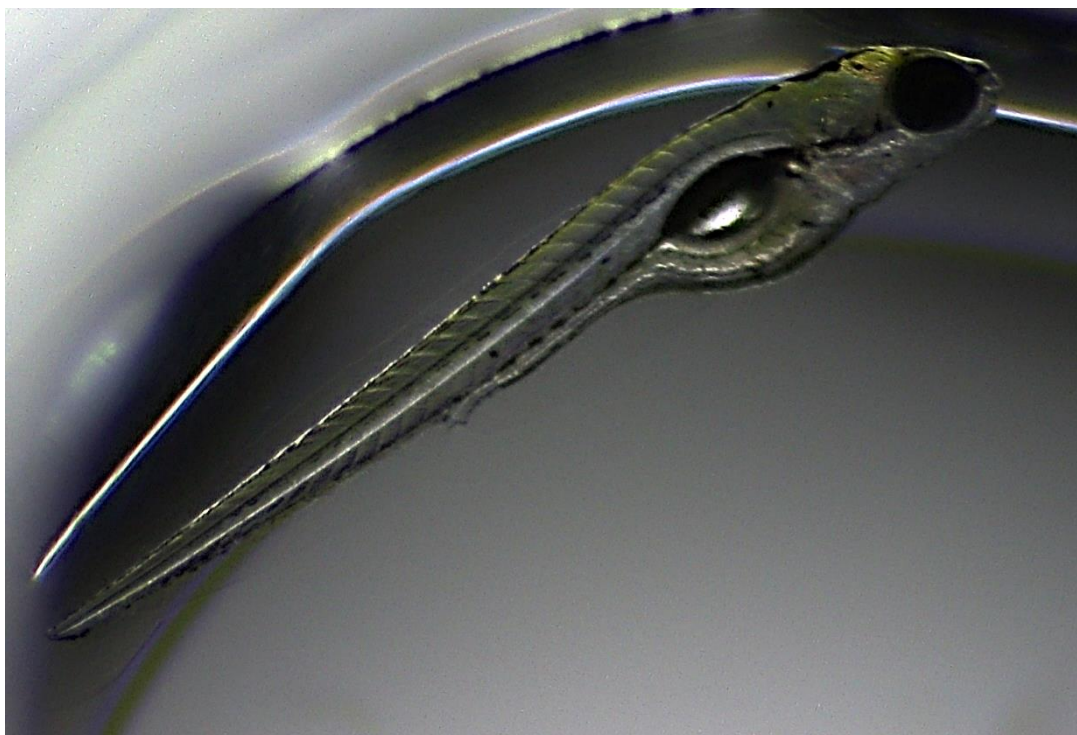
Figure 3.24: A fully developed *Danio rerio* zebrafish at five days post fertilization. Zebrafish embryos were grown in E3 media until five days post fertilization after which pictures were taken using an Accu-Scope 1906420 microscope using an Excelis HDS Lite camera and monitor system (Excelis, USA).

The nine endpoints were first examined after exposure to DMSO to create internal points of reference specific for the zebrafish used for this study, due to the nature of their origin. The somites present in five-day old zebrafish were observed to be clearly defined along the length of the tail, on either side of the notochord as displayed in Figure 3.25 A. These clear lines were seen to splay out from the centre of the tail as was expected from correctly formed somites. The notochord was seen to extend from the frontal region of the swim bladder to the tip of the tail, being clearly delineated from the somites along the entire length of the tail as depicted in Figure 3.26 A. The tail was seen to be straight, without any kinks or protrusions, as demonstrated in Figure 3.27 A. All five fins were seen to be correctly developed, with clear smooth edges on each of the dorsal fin, along the top of the tail, the ventral fin underneath and the caudal fin along the point of the tail, as indicated in Figure 3.28 A. Each pectoral fin on either side of the body was also seen to possess the correct shape (Figure 3.28 B).
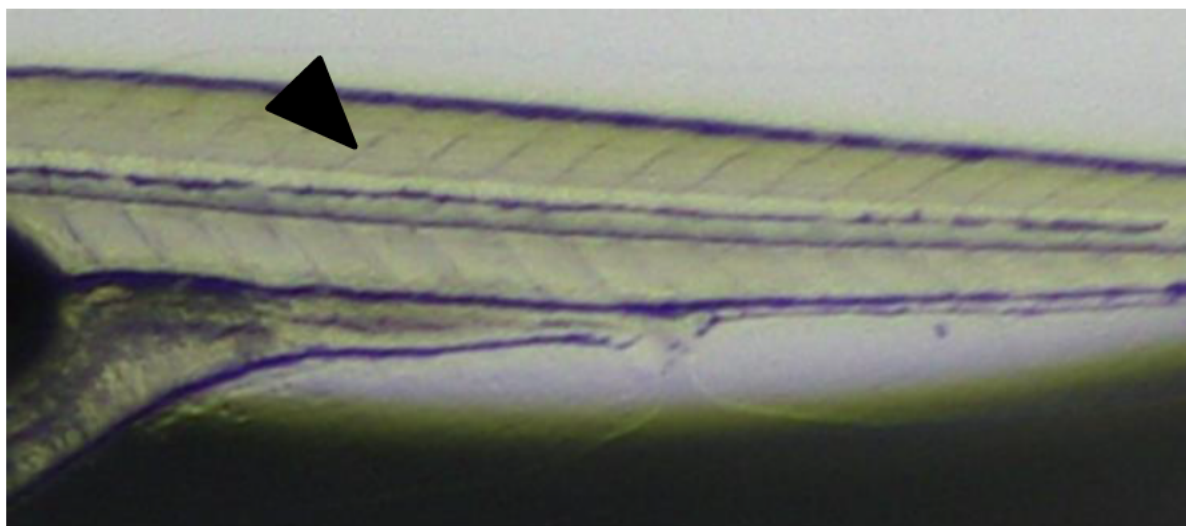
Figure 3.25: Somite morphologies in the tails of *Danio rerio* zebrafish at five days post fertilization after exposure to various compounds. The somite structure in the tail of a zebrafish after exposure to DMSO only, in panel A with the arrow indicating the clearly defined line that delineates a somite from adjacent somites. Panel B depicts somites resulting from exposure to 128.6 μM **C34**, allocated a somite morphology score of two, with the encircled area demonstrating a region with particularly poor somite definition.

For the facial structure, the otic capsule was observed to be the correct distance from the eye, with the mesenchyme region found between the otic capsule and the eye. Furthermore, the olfactory region was seen to be fully developed and at the expected size indicating a correctly formed facial region as indicated in Figure 3.29 A. The lower jaw was also confirmed to be in the correct size and position and the pharyngeal arches were seen as clearly defined diagonal lines running along the base of the head indicating correct formation of the jaw and pharynx respectively. Figure 3.30 A depicts the profile of the zebrafish head, with the three segments of

the brain clearly defined, demonstrating the brain had formed correctly. Finally, the morphology of the heart was observed with the heart being present in between two indents in the base of the fish body. Both the atrial and ventricular chambers were seen to be present in the correct proportions as well as evidenced by Figure 3.31 A. All these regions were seen to be correctly formed in accordance with current literature attesting to the viability of the zebrafish models use as a means to measure morphological aberrations (Panzica-Kelly *et al.*, 2010).

A
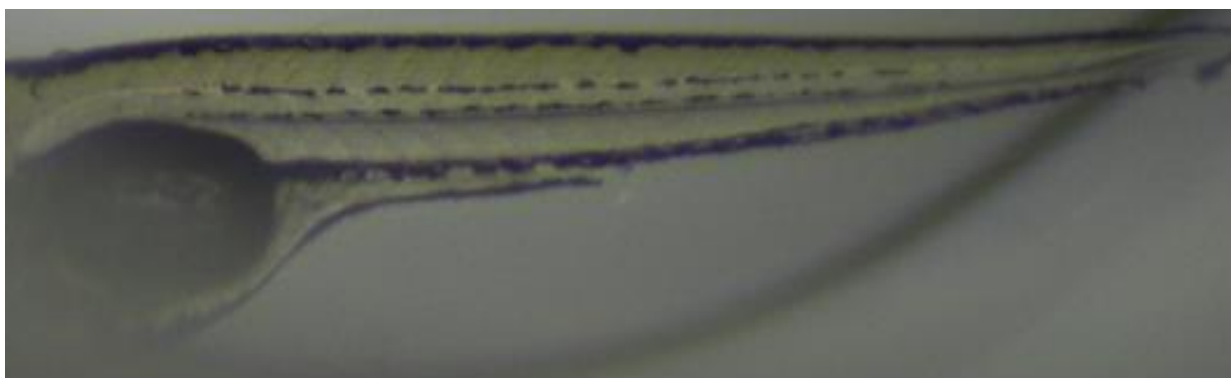
B

Figure 3.26: Notochord morphologies in the tails of *Danio rerio* zebrafish at five days post fertilization after exposure to various compounds. Panel A presents the full tail of a zebrafish after exposure to DMSO only, with the notochord running along the full length of the tail. Panel B displays a zebrafish exposed to **C34** at a concentration of 128.6 µM, exhibiting a notochord scoring a three on the morphology scale due to a large decrease in definition with the encircled area showing a severe lack of notochord definition.

3.9.3. *Determining compound teratogenicity on each morphological endpoint of Danio rerio* **embryos**

After establishing the validity of the zebrafish model, test compounds were assayed in an effort to determine any potential teratogenicity. For **C19**, only exposure at the highest non-lethal

concentration of 36.5 µM was seen to result in statistically significant morphological changes in the embryos (p-value < 0.05). Compound **C34** however exhibited teratogenicity at a range of non-lethal concentrations with morphology scores being statistically lower than those of the DMSO control at concentrations above 85.7 µM. Exposure to compound **C34** resulted in a significant effect on the development of eight of the nine structures being examined in Figure 3.32. No significant teratogenicity was detected in **C39** within the non-lethal concentration range (p-value < 0.05). Similarly to the lethality data, no significant teratogenicity was observed within the compound's solubility range. The morphological scores alongside lethality of each compound are depicted in Figure 3.33.

At all tested concentrations, **C34** resulted in a significant decrease in the morphological score of somites resulting in severe lack of definition in the somites of five-day old individuals, (Figure 3.25 B). The notochords of five-day old embryos exposed to 107.1 and 128.6 µM of **C34** developed with regions that were poorly defined as depicted in Figure 3.26 B. The tail was the only morphological endpoint that exhibited no significant decrease in integrity as compared to the control however several individuals were seen to possess mild kinks in their tails, as presented in Figure 3.27 B. Individuals exposed to 107.1 µM **C34** developed incorrectly formed pectoral fins, however this was not seen in individuals exposed to 128.6 µM **C34**. This is potentially the result of the effect that **C34** had on the swim bladders of the fish. The five-day old fish exposed to 128.6 µM **C34** exhibited deflated swim bladders resulting in several of these fish being unable to swim upright, resulting in only the dorsal, ventral and caudal fins being clearly visible. The lack of clearly observable pectoral fins potentially resulted in incorrect scoring, indicating a weakness in this scoring method, which could suggest that **C34** may be partially more teratogenic than indicated at this concentration. The fish exhibiting poorly developed fins (Figure 3.28 C and D), possess bent, misshapen pectoral fins as well as dorsal and ventral fins with irregular edges. Notably the jaws, pharyngeal arches and facial scores were all significantly lower in individuals exposed to **C34** at concentrations of 85.7, 107.1 and 128.6 µM, exhibiting misshapen mesenchyme regions, incorrectly proportioned olfactory regions, misshapen lower jaws and indistinguishable pharyngeal arches, as depicted in (Figure 3.29 B). The brain was also affected by the same concentrations as the facial structures, particularly with individuals displaying reduced forebrains as compared to those in the solvent control, displayed in Figure 3.30 B. Finally, the hearts of individuals exposed to **C34** at all non-lethal concentrations, from 64.3 to 129.6 µM, were significantly malformed, with the most incorrectly devloped hearts being severely enlarged and misshapen, presenting alongside severe

oedema around the heart (Figure 3.31 B). It is notable that all of the structures assessed, that are present in the head, were affected by the same concentrations of **C34**, namely the brain, jaw, face and pharyngeal arches indicating that **C34** may affect these structures through the same mechanism of action. It was also noted that only the heart and somites were affected by all test concentrations. The latter being of interest as, although the somites were less defined at all concentrations, no significant decrease in the scoring of the tail occurred at any concentration, despite the somites forming a major part of the tail.

A
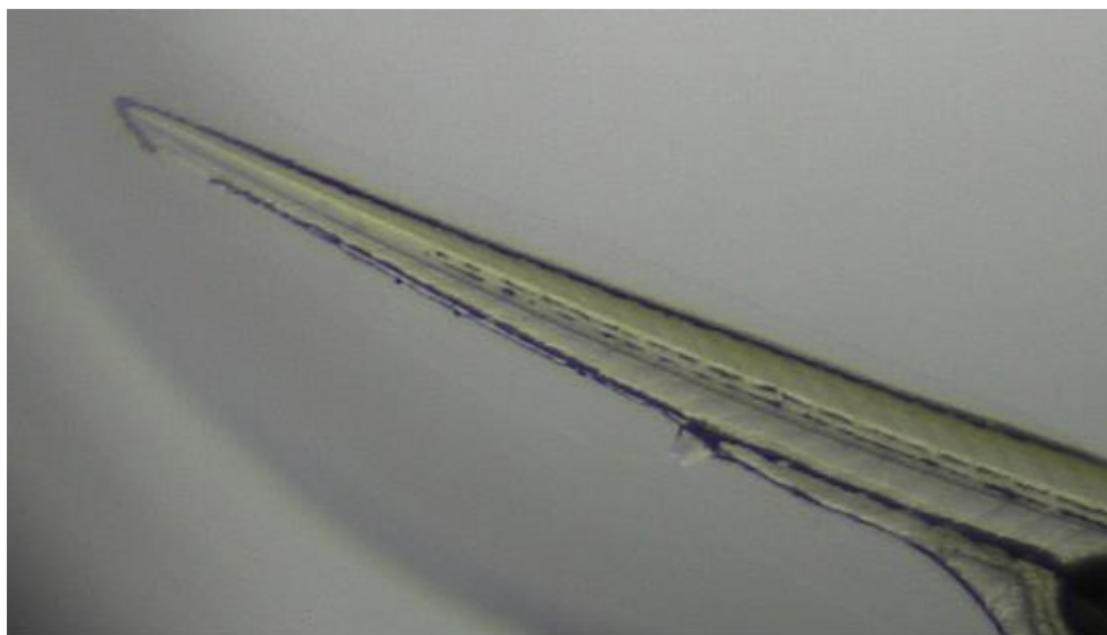
B
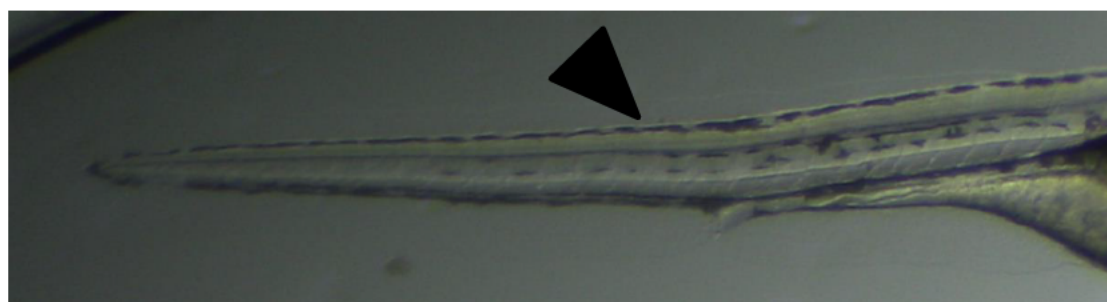
Figure 3.27: Tail morphologies in *Danio rerio* zebrafish at five days post fertilization after exposure to various compounds. The tail in panel A is an example of a tail from an embryo exposed to DMSO only. Panel B is an example from an individual exposed to 36.5 μM **C19** with an incorrectly formed tail with a kink, indicated by the arrow, which was the basis for the morphology score of four for this individual.
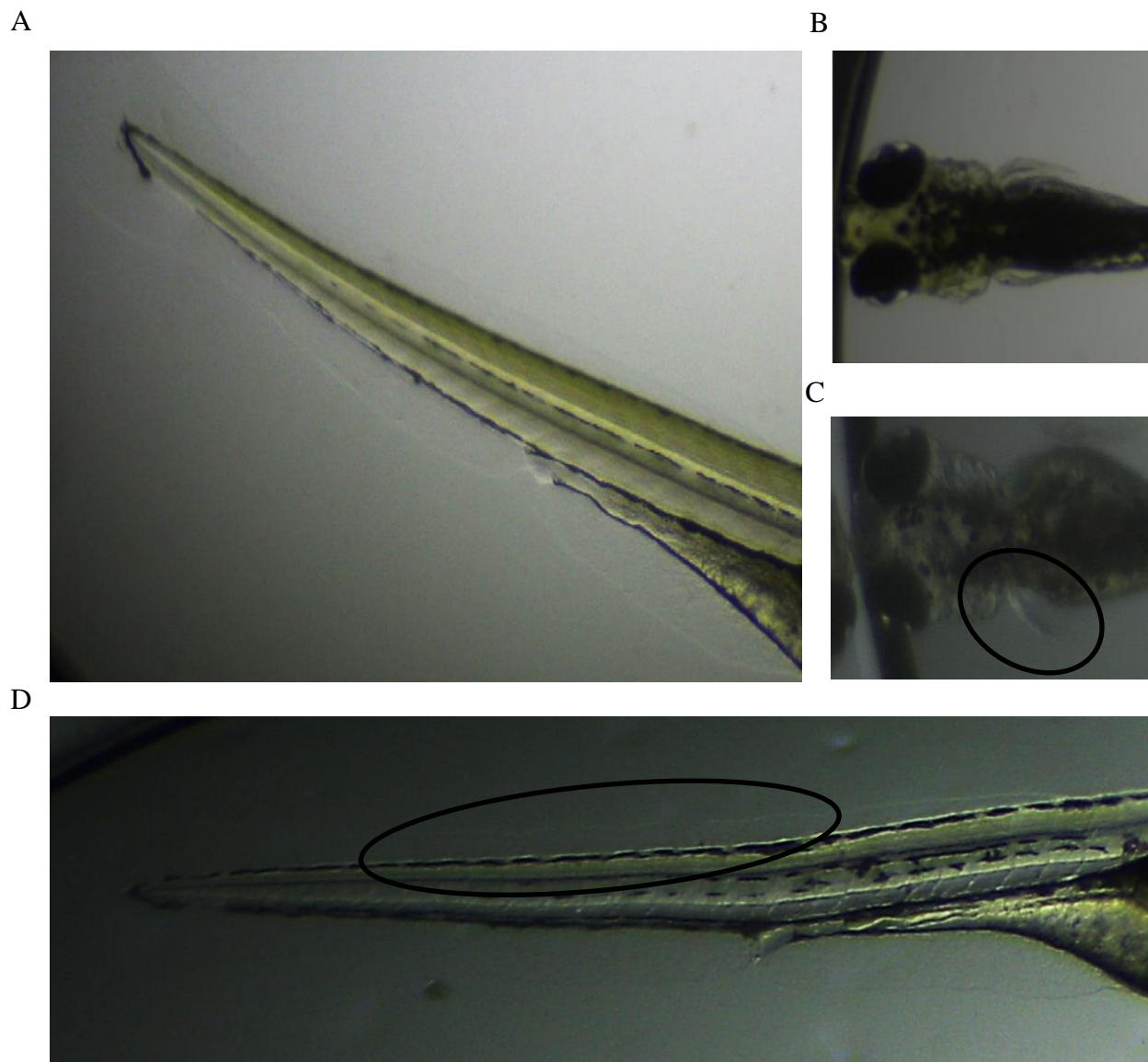
Figure 3.28: Fin morphologies in *Danio rerio* zebrafish at five days post fertilization after exposure to various compounds. The three tail fins from a zebrafish exposed to DMSO are depicted in panel A. Panel B presents the top view of the head of the fully developed zebrafish with the pectoral fins visible. Panel C is the top view of a zebrafish exposed to 128.6 µM **C34**, with a deformed left pectoral fin, which has been encircled, resulting in a score of three. Highlighted in panel D is a deformed dorsal fin from an individual exposed to 128.6 µM **C34**, demonstrating an irregular edge which has been encircled, resulting in a morphology score of three for fins in this case.
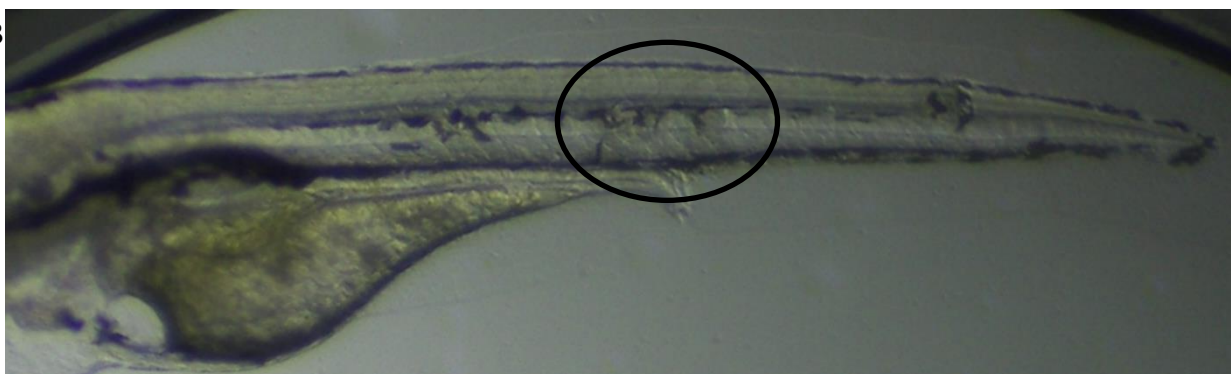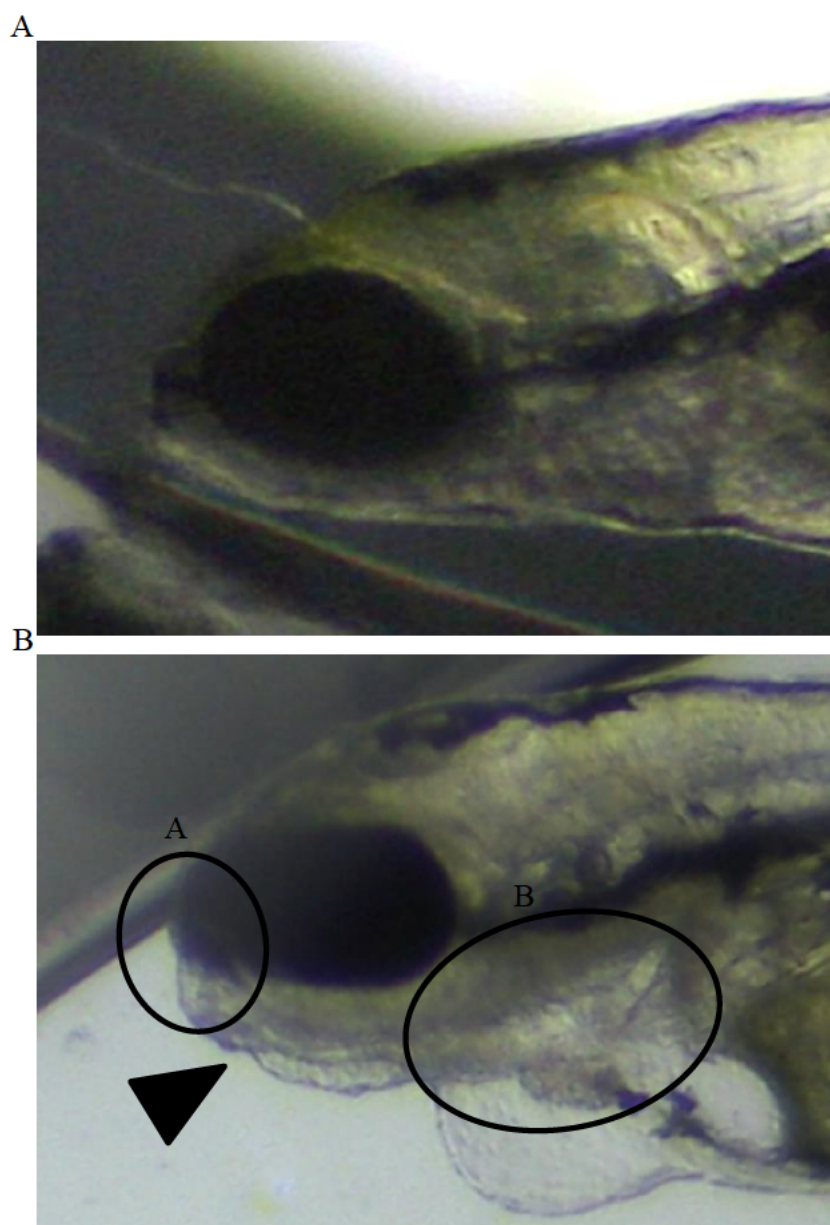
Figure 3.29: Facial structure, jaw and pharyngeal arch morphologies in *Danio rerio* zebrafish at five days post fertilization after exposure to various compounds. The fully developed head structure of a zebrafish, exposed to DMSO only, is depicted in panel A. Panel B presents an individual exposed to 128.6 μM **C34** exhibiting an underdeveloped olfactory region, encircled and labelled A, and a misshapen lower jaw, highlighted *via* the arrow, resulting in a facial morphology score of three and a jaw morphology score of three respectively. The zebrafish also has indistinguishable, poorly defined pharyngeal arches, encircled and labelled B, resulting in a pharyngeal arch morphology of two.
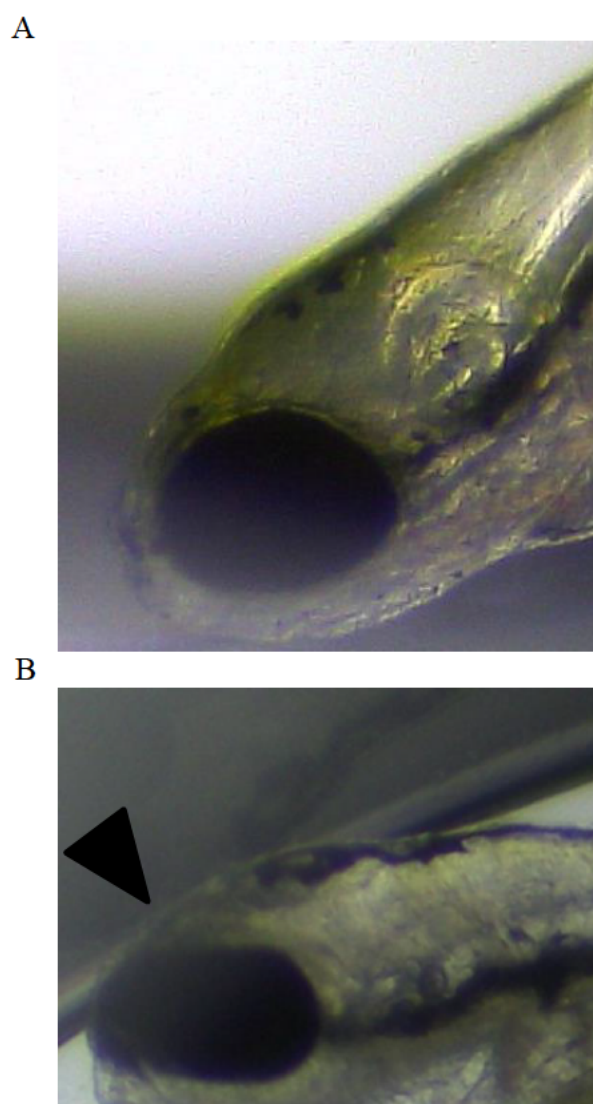
Figure 3.30: Brain morphologies in *Danio rerio* zebrafish at five days post fertilization after exposure to various compounds. The head of a zebrafish exposed to DMSO is depicted in panel A with the brain segmentation clearly seen. Panel B displays the head of an individual exposed to 128.6 μM **C34** exhibiting an under formed frontal brain segment, highlighted by the arrow, which was the basis of the score of three for the brain of this fish.

Whether a compound is classified as teratogenic is based on the compound's TI, which is the ratio of the compounds EC_{50} and LC_{50}. If this value is greater than one, the compound is deemed teratogenic (Alafiatayo *et al.*, 2019). Only **C34** yielded a sufficient decrease in morphological score to allow for the calculation of an EC_{50} value, however, indicating **C19**, **C39** and **C5064** were not teratogenic. Using the lethality data as well as the morphological scores obtained from exposing the zebrafish embryos to the test compounds, the LC_{50} and EC_{50} values were calculated. The resulting teratogenic index of 1.33 classified **C34** as teratogenic. Despite this

however, it was noted that concentrations of **C34** at which no significant decrease in total morphological score were observed, were still higher than the lethal doses of **C19** and **C39**.
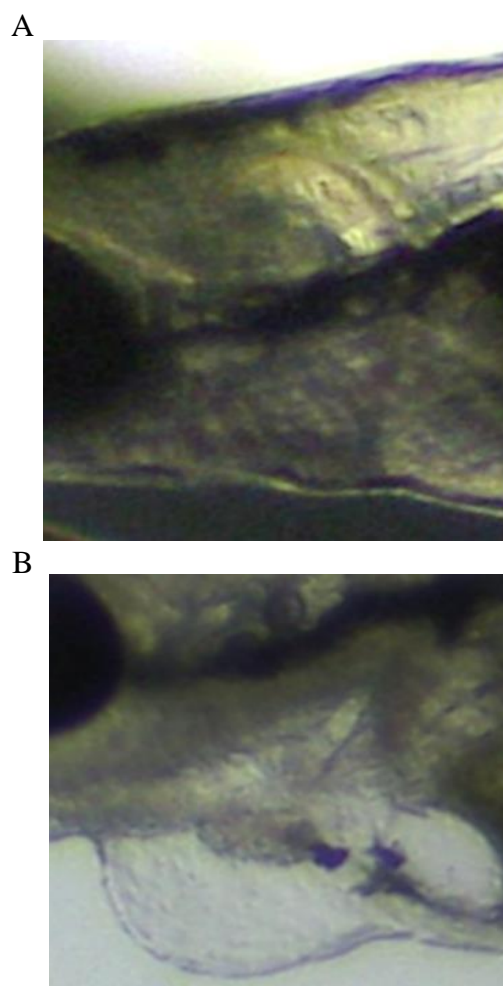
Figure 3.31: Heart morphologies in *Danio rerio* zebrafish at five days post fertilization after exposure to various compounds. The profile of the head of a zebrafish exposed to DMSO only centred on the heart. Panel B is an example of a zebrafish exposed to 128.6 µM **C34** with a heart that is misshapen and swollen, necessitating a score of two for heart structure.
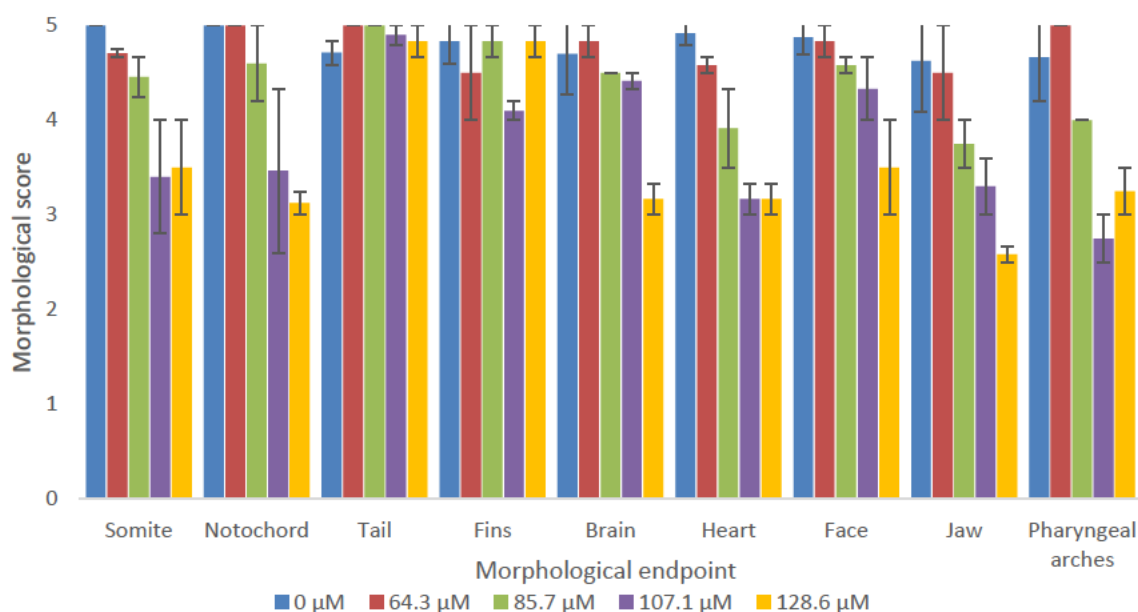
Figure 3.32: The morphology scoring for each morphological endpoint of *Danio rerio* embryos after exposure to compound C34 over a concentration range. Ten zebrafish per compound concentration were exposed to compounds at 3 hours post fertilization and incubated at 28.5°C for five days post fertilization (dpf) in E3 buffer. A scoring of correct development of the somites, notochord, tail, fins, heart, brain, facial structures, jaw and pharyngeal arches were determined after five days. Scoring for each morphological endpoint was out of five with a score of 5 indicating a correctly developed structure at five dpf. Error bars indicate the standard deviation of two replicates. The scoring for the control exposed to dimethyl sulfoxide is indicated in blue, red indicates the results after exposure to 64.3 μM **C34**, green represents the results from exposure to 85.7 μM, exposure to 107.1 μM is represented in purple and orange depicts the results after exposure of the embryos to 128.6 μM **C34**.
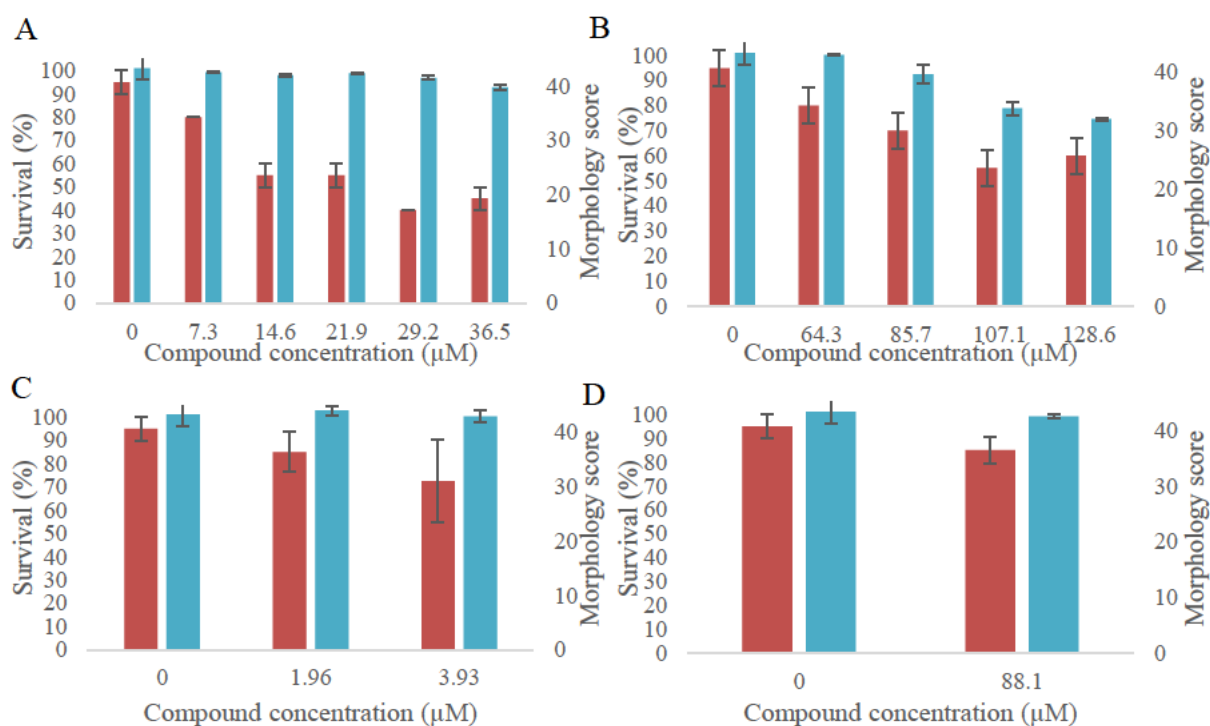
Figure 3.33: The survival rate and teratogenicity scoring of zebrafish exposed to various compounds at concentration ranges. Ten zebrafish per compound concentration were exposed to compounds at 3 hours post fertilization and incubated at 28.5°C for five days post fertilization (dpf) in E3 buffer. Survival and a scoring of correct development of the somites, notochord, tail, fins, heart, brain, facial structures, jaw and pharyngeal arches were determined after five days. Scoring for each morphological endpoint was out of five with a total score of 45 indicating a correctly developed zebrafish at five dpf. Error bars for survivability and teratogenicity indicate the standard deviation of two biological replicates (n = 2). The zebrafish embryo toxicity test results for compound **C19** are shown in A, **C34** displayed in B, **C39** in C and **C5064** in D. Red bars represent zebrafish survival at five dpf whereas blue bars indicate teratogenicity.

3.10. Utilizing in silico tools to predict the drug-likeness of test compounds

The in silico predictions for the drug likeness of the test compounds, as tabulated in Table 3.5, indicate that all test compounds are blood brain permeable, which is of particular importance if these compounds are to reach BACE1 in the brain. Compound **C5064**, however, was predicted as being susceptible to removal from the central nervous system by P-glycoprotein which is of interest should **C5064** see further use *in vivo* as a BACE1 inhibitor. None of the compounds were identified as pan assay interference compounds by SwissADME nor as aggregators by Aggregate advisor, which is important should any of the test compounds be used in activity assays. With regards to solubility, three scores were calculated for each compound predicting solubility in water however, it was noted that the Ali solubility prediction was the closest to the experimentally calculated solubility of **C5064**. The SILICOS-IT score underestimated the

solubility of **C19**, **C34** and **C5064** whereas the ESOL solubility score overestimated the solubility of **C5064**, predicting a solubility double that of the solubility seen in this study. The OSIRIS property explorer predicted that only **C39** was mutagenic or had a reproductive effect, which was of interest given the high levels of toxicity observed for **C39** in the ZFET. It was also noted that the drug score provided by OSIRIS suggests **C39** to be the least drug like, with **C34** being the most drug like of the initial test compounds. Compound **C5064** produced the highest drug score of 0.82.

Table 3.5: The predicted drug-likeness of the test compounds using in silico SwissADME, Aggregate advisor and OSIRIS analysis tools.

Category	Measurement	Compound			
		C19	C34	C39	C5064
Absorption	GI absorption[a]	High	High	High	High
	BBB permeant[b]	Yes	Yes	Yes	Yes
	PGP substrate[c]	No	No	No	Yes
Pharmacokinetics	Cytochromes inhibited[d]	1	2	0	3
Drug likeness	Lipinski[e]	0	0	0	0
	Bioavailability score[f]	0.55	0.55	0.55	0.55
	PAINS[g]	0	0	0	0
Lipophilicity	Consensus log $P_{o/w}$[h]	2.344	2.32	1.788	2.902
Solubility	ESOL Solubility[i] (mg/ml)	0.158	0.471	0.603	0.0764
	Ali Solubility[j] (mg/ml)	0.365	0.232	0.976	0.0268
	SILICOS-IT Solubility[k] (mg/ml)	0.00504	0.0092	0.083	0.00212
Aggregate Advisor	Aggregator[l]	No	No	No	No
OSIRIS predictions[m]	Mutagenic	No	No	High risk	No
	Tumorigenic	No	No	No	No
	Irritant	No	No	No	No
	Reproductive effect	No	No	Medium risk	No
	Drug score[n]	0.49	0.66	0.44	0.82

[a] Gastrointestinal absorption

[b] Blood brain barrier permeation

[c] P-glycoprotein substrate

[d] Predicted amount of P450 cytochrome enzymes the compound is capable of inhibiting; the specific cytochromes tested being P450 CYP1A2, P450 CYP2**C19**, P450 CYP2C9, P450 CYP2D6 and P450 CYP3A4

[e] Lipinski's rule of five filter (Lipinski *et al.*, 2001)

f   Abbot bioavailability score (Martin, 2005)

g   Pan assay interference structures (Baell and Holloway, 2010)

h   Consensus Log $P_{o/w}$ calculated from iLOGP, XLOGP3, WLOGP, MLOGP and SILICOS-IT lipophilicity scores (Daina *et al*., 2014, Lipinski *et al*., 2001, Wildman and Crippen, 1999)

i   Topological solubility score (Delaney, 2004)

j   Topological solubility score (Ali *et al*., 2012)

k   Fragment-based solubility score calculated by FILTER-IT program version 1.02

l   http://advisor.bkslab.org/ (Irwin *et al*., 2015)

m Predictions made utilizing the OSIRIS property explorer

n   OSIRIS score calculated based off of predicted scores for cLogP, logS, molecular weight, mutagenicity, tumorigenicity, irritant and reproductive effect scorings

SwissADME accessed*via*http://www.swissadme.ch/ last accessed in December 2020

OSIRIS property explorer accessed*via*https://www.organic-chemistry.org/prog/peo/ last accessed in December 2020

CHAPTER 4: DISCUSSION

It was the goal of this study to determine if utilizing BACE1 thermostability as a means of detecting potential binding molecules was an effective strategy. To accomplish this, first, the conditions for the expression of recombinant BACE1 were explored in an effort to determine possible rBACE1 yields. Secondly, the produced rBACE1 was assessed for its usability in a compound screening assay utilizing its thermostability as a means to detect protein-ligand binding. Thirdly, an in silico BACE1 ligand screening assay was compared to the assay successfully developed in objective two to determine its reliability in later stages of the study. Fourth, the utility of intracellular BACE1 thermostability as a means to determine intracellular protein-ligand binding was assessed using the test compounds successfully detected in objective two. Fifth, the use of the test compound capable of intracellular BACE1 binding was investigated for its potential as a lead compound for development of further, more potent, BACE1 binding ligands. Finally, any *in vivo* toxicity of the test compounds was examined, to assess any potential clinical viability.

4.1. Greater recombinantly expressed BACE1 yields achieved through optimization of expression conditions

A truncated form of the human BACE1 gene was synthesized in a pGEX4T-1 vector. *E. coli* BL21 (DE3) cells were then transformed with this plasmid. The purity of the plasmid was assessed using restriction digestion; however, anomalous bands were present in both the undigested and digested plasmid samples isolated from the transformed *E. coli*. This indicated that there was potentially contaminating plasmid DNA within the *E. coli* cells and as a result further confirmation of the presence of the correct BACE1 gene was carried out through sequencing with the identical sequence to the original GenBank sequence (GeneID: 23621) being confirmed to be present in the *E. coli* cells. Transformed *E. coli* cultures were grown, and expression was induced using IPTG. Expression of rBACE1 was compared under various conditions, where expression in LB media yielded the highest concentration of soluble rBACE1-GST. It was anticipated that LB media would yield less rBACE1-GST compared to 2xYT due to less nutrients. The lower overall protein yield could thus result in less aggregation and thus less inclusion bodies, resulting in more soluble fraction. In contrast, the largest yields of insoluble rBACE1-GST were generated after induction using 0.5 mM IPTG or following induction in the presence of 1% ethanol using 1 mM IPTG. Due to this, a combination of the two conditions, namely induction with 0.5 mM IPTG in the presence of 1% ethanol, was tested.

It was found however, that 16°C overnight induction in the presence of 1 mM IPTG and 1% ethanol was found to result in optimal expression. Initially 1% ethanol was added to increase the solubility of the expressed rBACE1-GST fusion protein however it was determined that this had the opposite effect. The addition of ethanol during induction is meant to slow down bacterial growth, resulting in less protein production and thus less inclusion body formation. The addition of 3% ethanol did decrease the amount of insoluble protein while increasing the amount of soluble protein indicating that 3% ethanol was sufficient in curbing protein production, however this evidently did not result in a complete reduction in inclusion body formation due to the presence of the majority of rBACE1-GST in the insoluble fraction.

It was noteworthy that the combination of 0.5 mM IPTG in conjunction with 1% ethanol did result in more soluble rBACE1-GST indicating that reducing protein production through lowering the amount of induction agent did affect inclusion body formation. The increase in insoluble protein as a result of the addition of 1% ethanol could be an indication that *E. coli* growth overnight might mitigate the growth retardation brought about by 1% ethanol and without the decreased IPTG, inclusion body formation was unaffected. This condition was pursued as it was observed that the insoluble fraction that resulted from this induction contained the largest amount of protein compared to any other fraction. These expression conditions were done at the cost of incorporating a solubilization and refolding step. Expression of the rBACE1-GST fusion protein was confirmed with both anti-GST IgY as well as anti-BACE1 IgG antibodies. Initial western blot analysis indicated degradation of the fusion protein over the course of expression, however, after successful affinity purification of the rBACE1-GST protein, this degradation was not observed. It was noted that 100% purity was not achieved as established through the detection of residual GST-tag with anti-GST antibodies. As this protein was not detected by Coomassie stain, it was determined that the residual GST was present at markedly low concentrations, thus indicating high purity of the desired fusion protein.

Both solubilization and refolding of rBACE1-GST were successfully carried out using a 10% sarkosyl solution. Other refolding techniques were attempted; however, they were demonstrated to be ineffective. Refolding was confirmed *via* the presence of protease activity in a gelatine gel zymogram. Gelatine gel zymography has not previously been described for BACE1 yet provides an inexpensive method to confirm activity. It also allows for the detection of dimer formation. The commercially purchased BACE1 produced in HEK-293 cells was tested as a positive control to confirm that BACE1 is indeed capable of gelatine degradation. It

was noted that the mammalian BACE1 exhibited activity in what was assumed to be a dimeric and tetrameric form; however, no activity was detected as a monomer. This is in contrast to rBACE1 which only displayed protease activity in its monomeric form. This is in accordance with literature as it has been reported that both the cytosolic and transmembrane domains assist in dimerization whilst the rBACE1 does not possess these domains (Schmechel *et al*., 2004). It was noted that this method was not viable for screening as the protease activity was unaffected by the presence of the known BACE1 inhibitor, Verubecestat. This could potentially be due to the Verubecestat being unable to readily access the BACE1 in the zymogram, however this has not been reported previously. The activity was not accurately quantifiable either, however it was noted that 1 µg of mammalian BACE1 resulted in a distinct band of proteolytic activity, whereas 100-fold more rBACE1 was required to achieve a similar level of activity. This is likely due to the effect of post-translational modifications on BACE1 proteolytic activity, as it has been demonstrated previously that the glycosylations found on mammalian BACE1 are required for maximal activity (Huse *et al*., 2000). It was also further noted that prior to the removal of the GST-tag from the recombinant protein, no activity was present indicating the tag disrupts access to the rBACE1 active site. As a result, cleavage of the GST-tag was carried out prior to all subsequent experiments. This expression and refolding protocol resulted in an overall rBACE1 yield 70% higher than the highest reported yield (Tomasselli *et al*., 2008). Most notably, it only requires a 24-hour refolding step as opposed to the 3-4 day refolding steps previously described (Tomasselli *et al*., 2008). Finally, the rBACE1 produced was proteolytically active.

A BACE1 expression system was also generated using the mammalian HEK-293 cells. These HEK-293 cells were stably transfected with a pcDNA-DEST 40 vector containing the full human BACE1 gene with transfection being confirmed by the ability of the HEK-293 cells to grow in the presence of G418-containing selection media. A non-truncated form of BACE1 was expressed in these mammalian cells as the transmembrane and cytoplasmic domains are required for transport of BACE1 to the Golgi apparatus. This was done to ensure proper localization of BACE1 inside the cells to the lysozomes. It was demonstrated through western blot that BACE1 was produced in both non-transfected and transfected variants of the HEK-293 cells which was expected as BACE1 expression in HEK-293 cells has been reported in literature previously (Vetrivel *et al*., 2011). Transfection with the BACE1-pcDNA-DEST 40 resulted in a higher expression of the BACE1 protein, with densitometry indicating 60% higher production of the 67 kDa BACE1 protein. Two large bands were present in all lysates

corresponding to sizes of approximately 30 and 35 kDa respectively, which were detected by the anti-BACE1 antibodies.

The rBACE1-GST was confirmed to be antigenic and capable of producing both anti-GST and anti-BACE1 antibodies in chickens. The anti-BACE1 IgY antibodies were purified and confirmed to bind to both bacterial and mammalian forms of BACE1. Specificity of both the anti-GST and anti-BACE1 antibodies were confirmed through western blot analysis. The ability to generate both anti-GST and anti-BACE1 antibodies from the same fusion protein indicates that neither the GST-tag, nor the BACE1 protein, hinder the immunogenicity of the other. Further, the generation of antibodies indicates that the fusion protein is immunogenic without the need for the attachment of carrier proteins, likely due to the size of the fusion protein. It was however noted that while the concentration of anti-BACE and anti-GST antibodies present in the eggs of the immunized chicken increased with each successive immunization, after the final immunization, the antibody concentration decreased and then plateaued. The BACE1 produced in *E. coli* cells presented a far less defined band when run on a western blot as compared to the HEK-293 BACE1. It was expected that the lack of glycosylations would allow for more binding of SDS, which would translate to clearer bands as compared to the glycosylated BACE1 from the HEK-293 cells however this did not occur. Furthermore, western blots of the purified rBACE1 exhibited smearing at the top of the lane indicating aggregation.

4.2. Bacterially produced rBACE1 is suitable for compound screening

In order to accomplish Objective 2 of the study, the potential for rBACE1 thermostability to be used as a means to screen for BACE1 ligands needed to be established. This was attempted using an in-house library of compounds in a DSF assay using the purified rBACE1. Initially, multiple buffers were used in an effort to determine the optimal conditions for the assay with PBS at a pH of 7.4 ultimately being selected. The lack of a discernible peak when using AMT buffers at the same pH as PBS suggested that either the AMT buffer itself interfered with the assay, or more likely, the buffer transfer resulted in protein denaturation. Interference in thermostability assays by buffer components has been previously reported (Long and Yang, 2009), however based on the lack of a defined peak in acetate buffer as well, it is likely that the buffer transfer denatured the protein. Since the purified protein was ultimately purified in PBS, using PBS as the assay buffer did not require a buffer transfer step as opposed to the use of the acetate and AMT buffers. Since aggregation has been reported as a potential side-effect of buffer transfers, which would prevent a peak from forming during the assay, this is likely

what occurred (Bondos and Bicknell, 2002). The pH at which the peak was detected, pH 7.4, was of interest as a study by Hayley *et al.* in 2009 demonstrated that BACE1 remains folded at pH 7.4 while becoming proteolytically inactive. It was also reported that this inactive form possesses a Tm value of approximately 84°C indicating thermostability exceeding physiological ranges. The study carried out by Hayley *et al.* however did not report a specific Tm value due to the limits of the circular dichroism technique used. This study successfully determined the Tm value of rBACE1 to be precisely 82°C indicating the DSF assay may potentially be better suited to studying proteins with abnormally high Tm values. The rBACE1, when assayed in PBS exhibited a single state unfolding, demonstrated by the R^2 value of the plot of $\Delta_u G°$ against temperature of > 0.9 (Wright *et al.*, 2017). This single state unfolding is in contrast to the findings of Hayley *et al.*, where a two-state transition was observed. This is potentially due to the use of a truncated protein as opposed to the full protein utilized by Hayley and co-workers. This suggests that the transition states, observed during BACE1 unfolding, involve the domains not present in the rBACE1 form used in this study. The thermodynamic parameters of rBACE1 unfolding in PBS, namely the $\Delta_u G°$, $\Delta_u S°$ and $\Delta_u H°$, were found to be 112 ± 0.1 kJ/mol, 1.82 ± 0.002 kJ/mol K and 642 ± 0.7 kJ/mol respectively. These values indicate high thermostability which was expected due to the high Tm value of 82 ± 0.21°C. Furthermore, these values possess low variances despite the extrapolation of the original data over such a large temperature range.

The DSF technique was confirmed to be an effective screening tool through a significant shift, with a p-value set to 0.05, in Tm of the recombinant protein in the presence of 1 μM of the specific BACE1 inhibitor, Verubecestat. This validated the use of the rBACE1 as a means to detect BACE1 binding compounds. Furthermore, the binding of Verubecestat confirmed that utilizing the inactive form of rBACE1 at pH 7.4 allowed for detection of catalytic site binding molecules. This binding was also confirmed to occur in the presence of 2% DMSO, however DMSO lowered the Tm values of BACE1 both in the presence and absence of Verubecestat. This was unsurprising as DMSO has been documented to lower reported Tm values in DSF experiments, with the maximum suggested DMSO concentration being 2% (Sorrell *et al.*, 2010). It was also noted that after the addition of either Verubecestat or DMSO, BACE1 exhibited unfolding characteristics implying the formation of a transition state. As a result, thermodynamic parameters could not be accurately predicted. This prevented the calculation of the change in $\Delta_u G°$ as a result of ligand binding indicating a limitation of this method of $\Delta_u G°$ calculation. After optimization of the DSF experiment, the 50-compound in-house library was

screened against the rBACE1 and three compounds produced significant Tm increases in rBACE1. These three compounds were, **C19** (N-ethyl-N',N'-dimethyl-N-[2-(trifluoromethyl) benzyl]-1,2-ethanediamine), **C34** (3-cyclopentyl-N-(4-pyridinylmethyl) propenamide) and **C39** (4-chloro-1-(2-ethoxybenzoyl)-1H-pyrazole). This led further credence to the validity of this approach of BACE1 ligand screening as it was capable of detecting novel BACE1 binding compounds.

4.3. In silico BACE1 ligand screening provides comparable findings to *in vitro* DSF screening

An in silico model of BACE1 was prepared and validated using the Schrodinger Biologics suite (2018-2). This was carried out as a means of determining the comparability between *in vitro* screening results from the DSF experiments and in silico docking experiments, as a reliable in silico BACE1 model would allow for screening of large compound libraries as a preliminary method of high throughput screening. The model was validated by successfully redocking Verubecestat into the BACE1 active site. This bound Verubecestat demonstrated a low RMSD when compared to the co-crystallized Verubecestat, further, many of the interactions between Verubecestat and BACE1 that are known to occur were successfully reproduced by this model. This model indicated that 5 of the 50 compounds bind to the BACE1 active site in a way that would be conducive to inhibition. Furthermore, two of the three compounds, namely **C19** (N-ethyl-N',N'-dimethyl-N-[2- (trifluoromethyl)benzyl]-1,2-ethanediamine) as well as **C34** (3-cyclopentyl-N-(4-pyridinylmethyl) propenamide), identified by the DSF ligand binding assay were identified as binders to the catalytic aspartate dyad of the BACE1 enzyme. The detection of both **C19** and **C34** highlighted as catalytic site binding ligands provided further merit to the theory that, whilst the BACE1 is in an inactive form in the DSF experiment, the active site remains intact.

The three compounds that were detected by this in silico model that did not bind in the DSF experiments are likely false positives. This infers that molecular dynamics simulations would be required as a means to filter out false positives. Nonetheless, as a means of high throughput screening as a preliminary step, the model is satisfactory as evidenced through the identification of both **C19** and **C34** as BACE1 ligands. Other potential reasons for the discrepancy between the DSF and docking results could be that these compounds do indeed bind the BACE1 catalytic site; however, either due to insolubility or aggregation, are incapable of binding to the protein in solution. This was not further explored however as this insolubility would have reduced their

potential as drug candidates as compared to the other compounds that were observed to bind in the DSF assay.

The lack of detection of **C39** as a BACE1 binding compound further necessitated a search for other binding sites present on the protein. Whilst docking of the three binding compounds to a detected non-catalytic binding site was also attempted, this investigation indicated that **C19** bound to this site in conjunction with the catalytic site whereas no indication of the binding mode of **C39** was revealed. Further study of the BACE1 protein surface would therefore be required, potentially into the only reported BACE1 allosteric site as binding to this site could potentiate the use of **C39** as a lead molecule. No binding was detected for **C34** to the non-catalytic site of BACE1 indicating that **C34** potentially binds specifically to the active site of BACE1. Docking against BACE2 was also carried out with only **C19** binding to the BACE2 active site. The lack of binding of **C34** to BACE2 reveals **C34** to be specific to BACE1, a feature that is valuable for human treatment. The binding of **C19** to the BACE2 active site is of some concern, not only to its potential as a BACE1 specific ligand, but also to the BACE1 model Since **C19** was detected to bind to all three sites that were tested in this study, it calls into question the accuracy of these additional binding site receptor grids. The likelihood of a compound binding specifically to two sites on the same protein as well as a site on another protein, despite the similarities in the proteins, is unlikely. The BACE1 catalytic site receptor did not warrant this concern as it was validated using Verubecestat. Validation was not possible for the non-catalytic BACE1 binding site detected using SiteMap however, due to the lack of known binding molecules. This lack of validation further necessitates the use of molecular dynamics simulations, should this binding site be further explored.

4.4. Detected compounds bind to BACE1 intracellularly

Compounds detected through DSF were assayed for the ability to bind BACE1 inside mammalian HEK-293 cells using the CETSA. Prior to carrying out the CETSA, the cytotoxicity of the test compounds was assayed to ensure toxicity would not interfere with the assay. No CC_{50} value was obtained for any of the test compounds within the concentration range that was used, confirming that compound toxicity would not be a confounding variable. Initially the T_{agg} of monomeric BACE1 in the presence of DMSO was determined to be $49.53 \pm 0.69°C$ in contrast to the T_m found for rBACE1 during DSF which was $82 \pm 0.21°C$. It was also noted that a CETSA could potentially be carried out using the dimer of BACE1 as this exhibited a clear temperature dependent decrease in concentration through the CETSA as well. A smaller

temperature range would need to be used however to determine the T_{agg} of the BACE1 dimer. The discrepancy between the two values of the monomeric BACE1, that both give inference to a protein's stability, is due to the difference in the properties that directly contribute to protein denaturation and aggregation respectively. The T_m value is based on protein denaturation which is determined by the disruption of intramolecular bonds such as internal hydrophobic interactions and disulphide bonds while a proteins T_{agg} is based largely on a protein's solubility, meaning full denaturation and disruption of all intracellular bonds is not required for aggregation to occur, and it thus requires comparatively less energy in the form of heat to induce aggregation than to fully denature a protein. This relationship between a protein's T_{agg} and Tm, where Tm is higher than the T_{agg} has been seen previously in granulocyte colony stimulating factor by Robinson and colleagues under several different conditions (Robinson *et al.*, 2018).

After the T_{agg} for BACE1 produced in HEK-293 cells was determined, the compounds detected as BACE1 binders through DSF, namely Verubecestat, **C19**, **C34** and **C39**, were assayed for intracellular binding through the CETSA. Both Verubecestat and **C34** resulted in a significant increase in the T_{agg} of BACE1 with Verubecestat resulting in an increase from $49.53 \pm 0.69°C$ to $53.27 \pm 0.89°C$ and **C34** inducing a shift to 53.09 ± 1.30. Both **C19** and **C39** conferred no significant change to BACE1 T_{agg} indicating that these compounds are likely not capable of reaching BACE1 inside the HEK-293. Verubecestat provided validation of this method of screening for intracellular BACE1 ligands. This assay is not amenable to high throughput however, solidifying the position of this assay late in the compound screening pipeline. Furthermore, it is of note that while the CETSA provides further confirmation that **C34** binds to BACE1, the CETSA is incapable of detecting inhibition. While assaying for BACE1 inhibition would be the intuitive next stage of this study, intracellular binding has been exploited through the use of PROTAC technology to reduce enzyme concentrations.

4.5. Compound C34 functions as a viable lead compound for detection of further intracellular BACE1 binding molecules

Due to the confirmed ability of **C34** to bind intracellular BACE1, this compound was used as the lead compound to screen the Enamine compound database for similar, more complex compounds capable of binding BACE1. An iterative screening approach for detection of similar compounds utilizing the in silico BACE1 model established in this study was utilized ultimately resulting in the detection of a compound exhibiting a Glide score of -10.121. This score was 65% lower than the Glide score of **C34**, -6.574. A CETSA was then carried out using this

compound, **C5064**, in an effort to determine if it possessed the ability to bind intracellular BACE1. Compound **C5064** was confirmed to bind intracellular BACE1 through this assay indicating the use of **C34** as a lead was successful. Based on the lower Glide score of **C5064** when compared to **C34**, **C5064** is expected to bind BACE1 with greater affinity than **C34**. This compound thereby provides merit to the workflow utilized in this study, validating that the use of thermostability as a means of detecting BACE1 ligands is effective.

4.6. Assessing potential for clinical viability of test compounds using *in vivo Danio rerio* toxicity models

The toxicity of the test compounds was assessed using a zebrafish model to assay for potential viability for use of the compounds as binders of BACE1 *in vivo*. The test compounds yielded an effect on the embryos prior to hatching, demonstrating passage through the chorion. This was anticipated due to their sizes being well below the suggested cut off of 3 kDa, as the OECD guidelines suggest studies involving compounds greater than 3 kDa require dechorionation to ensure compounds are capable of accessing the embryo. The toxicity of each compound was determined through the use of a ZFET, allowing for LC_{50} values to be calculated. It was observed that **C39** exhibited the highest toxicity levels of the test compounds, with an LC_{50} value six times less than that of **C19**, which was itself five times less than that of **C34**. With regards to the toxicity of **C39**, it was noted that the in silico analysis predicted the compound to be mutagenic. Disruption of DNA may thus be the reason toxicity was observed so early in the development of the embryos exposed to **C39**. It was seen that the LC_{50} value of **C39** was lower than 3,4-DCA, a compound utilized as a positive control for toxicity tests due to its high toxicity with **C19** showing an LC_{50} value only 25% greater. These low LC_{50} values in combination with their confirmed inability to bind intracellular BACE1, reiterate the lack of viability of both **C39** and **C19** for further study. The LC_{50} value of **C34** was acceptable and provided justification for further study into its ability to bind BACE1 *in vivo*. No LC_{50} value was calculated for **C5064** due to limited solubility in the assay buffer however it was seen that **C5064** was incapable of inducing a toxic effect on the embryos within the limits of its solubility. This could be resolved through the dissolution of **C5064** in a less toxic organic solvent than DMSO, such as polyethylene glycol-400 as potentially this would allow greater concentrations of **C5064** to be used without the toxic effects of the solvent interfering with the results (Maes *et al*., 2012). No TI could be calculated for any of the other test-compounds indicating that while **C19** and **C39** are highly toxic, they are not classified as teratogens. The teratogenicity of

C34 was only observable at concentrations higher than the lethal doses of **C19** and **C39** however, indicating the potential for further use of **C34** as a means to bind BACE1 *in vivo*.

4.7. Conclusion

A recombinant form of the BACE1 protease was successfully expressed at yields higher than those currently found in literature. The protein was confirmed to be in an active conformation through the use of a gelatine gel zymogram, a technique which, through this study, has now been demonstrated to be a reliable method to confirm BACE1 activity. This pure, active rBACE1 was then used to create and optimize a DSF ligand binding assay, which was effectively used to screen an in-house compound library. This assay confirmed the use of the thermostability of BACE1 as a viable means of screening for BACE1 binding compounds as a total of three compounds were detected to bind to the recombinant protein through this assay. Molecular modelling was used to dock these ligands in silico, against a validated model of BACE1, whereby it was indicated that two of the three compounds bound to the active site, whereas the third compound possibly bound to an allosteric site. The recombinant BACE1 protein was also capable of inducing the production of antibodies in chickens, which were capable of detecting mammalian BACE1, expressed in a transfected HEK-293 cell line.

A CETSA to detect BACE1 ligands was successfully carried out and validated through the confirmation of binding by Verubecestat, with **C34** being demonstrated to bind BACE1 intracellularly. Compound **C34** was used as a lead compound to screen the Enamine Real database of compounds using the in silico model of BACE1 established in this study. Through this method **C5064** was identified as a potential BACE1 binding compound. The ability of **C5064** to bind BACE1 intracellularly was then confirmed using the CETSA once again with this discovery providing further validation of the successful establishment of a pipeline for screening for BACE1 ligands primarily utilizing the thermostability of BACE1 as a means to detect ligand interactions. This successfully confirmed the hypothesis that the CETSA could indeed be used as a tool to identify BACE1 ligands. This also fulfilled the aim of the project, of thoroughly assessing the thermostability of BACE1, through the determination of the Tm, T_{agg}, standard Gibbs free energy, entropy and enthalpy of unfolding as well as identifying that truncated BACE1 undergoes single state unfolding when exposed to temperature.

4.8. Further Studies

A FRET assay could be made and used to assess if **C19**, **C34** and **C39** can inhibit BACE1 activity. Confirmed inhibitors could then be optimized using QSAR to improve their inhibitory capabilities. Binding modes of these compounds could also be confirmed using X-ray crystallography. With regards to compound **C5064**, further study could be carried out to ascertain its toxicity and teratogenicity in the zebrafish model. Furthermore, **C5064** and **C34** could be used as leads to screen for more similar compounds in an attempt to identify more intracellular BACE1 binding molecules, potentially with greater solubility than that of **C5064**. Finally, the development of PROTACs using **C34** and **C5064** could be explored as possible means to target intracellular BACE1.

CHAPTER 5: REFERENCES

AIT-BOUZIAD, N., LV, G., MAHUL-MELLIER, A.-L., XIAO, S., ZORLUDEMIR, G., ELIEZER, D., WALZ, T. & LASHUEL, H. A. 2017. Discovery and characterization of stable and toxic Tau/phospholipid oligomeric complexes. *Nature Communications,* 8**,** 1678.

ALAFIATAYO, A. A., LAI, K. S., SYAHIDA, A., MAHMOOD, M. & SHAHARUDDIN, N. A. 2019. Phytochemical evaluation, embryotoxicity, and teratogenic effects of *Curcuma longa* extract on zebrafish (*Danio rerio*). *Evidence-Based Complementary and Alternative Medicine,* 2019.

ALI, J., CAMILLERI, P., BROWN, M. B., HUTT, A. J. & KIRTON, S. B. 2012. In silico prediction of aqueous solubility using simple QSPR models: the importance of phenol and phenol-like moieties. *Journal of Chemical Information and Modeling,* 52**,** 2950-2957.

ALMQVIST, H., AXELSSON, H., JAFARI, R., DAN, C., MATEUS, A., HARALDSSON, M., LARSSON, A., MOLINA, D. M., ARTURSSON, P., LUNDBÄCK, T. & NORDLUND, P. 2016. CETSA screening identifies known and novel thymidylate synthase inhibitors and slow intracellular activation of 5-fluorouracil. *Nat. Commun.,* 7**,** 11040.

ALZHEIMER, A. 1907. Uber eine eigenartige erkrankung der hirnrinde. *Allgemeine Zeitschrift fur Psychiatrie und phychish-Gerichtliche Medizin,* 64**,** 146-148.

ALZHEIMER, A. 1911. Über eigenartige Krankheitsfälle des späteren Alters. *Zeitschrift für die Gesamte Neurologie und Psychiatrie,* 4**,** 356-385.

BAELL, J. B. & HOLLOWAY, G. A. 2010. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *Journal of Medicinal Chemistry,* 53**,** 2719-2740.

BIN, Y., ZHU, Q., LI, M. & XIA, J. 2019. Comprehensive Analysis of Alzheimer's Disease Biologically Candidate Causal Genes Revealed by Function Association Study With GWAS. *IEEE Access,* 7**,** 114236-114245.

BIOGEN 2019. Biogen annual report 2019.

BOHRMANN, B., BAUMANN, K., BENZ, J., GERBER, F., HUBER, W., KNOFLACH, F., MESSER, J., OROSZLAN, K., RAUCHENBERGER, R., RICHTER, W. F., ROTHE, C., URBAN, M., BARDROFF, M., WINTER, M., NORDSTEDT, C. & LOETSCHER, H. 2012. Gantenerumab: A Novel Human Anti-Aβ Antibody Demonstrates Sustained Cerebral Amyloid-β Binding and Elicits Cell-Mediated Removal of Human Amyloid-β. *Journal of Alzheimer's Disease,* 28**,** 49-69.

BONDOS, S. E. & BICKNELL, A. 2002. Detection and prevention of protein aggregation before, during, and after purification. *Analytical Biochemistry,* 316**,** 223-231.

BRADFORD, M. M. 1975. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding *Analytical Biochemistry,* 72**,** 248-254.

BURGESS-BROWN, N. A., SHARMA, S., SOBOTT, F., LOENARJ, C., OPPERMANN, U. & GILEADI, O. 2008. Codon optimization can improve expression of human genes in Escherichia coli: A multi-gene study. *Protein Expression And Purification,* 59**,** 94-102.

BURKI, T. 2018. Alzheimer's disease research: the future of BACE inhibitors. *The Lancet,* 391**,** 2486.

CAI, H., WANG, Y., MCCARTHY, D., WEN, H., BORCHELT, D. R., PRICE, D. L. & WONG, P. C. 2001. BACE1 is the major β-secretase for generation of Aβ peptides by neurons. *Nature Neuroscience,* 4**,** 233-234.

CHANG, V. T., CRISPIN, M., ARICESCU, A. R., HARVEY, D. J., NETTLESHIP, J. E., FENNELLY, J. A., YU, C., BOLES, K. S., EVANS, E. J., STUART, D. I., DWEK, R. A., JONES, E. Y., OWENS, R. J. & DAVIS, S. J. 2007. Glycoprotein structural genomics: solving the glycosylation problem. *Structure,* 15**,** 267-273.

CHERET, C., WILLEM, M., FRICKER, F. R., WENDE, H., WULF-GOLDENBERG, A., TAHIROVIC, S., NAVE, K. A., SAFTIG, P., HAASS, C., GARRAT, A. N.,

BENNETT, D. L. & BIRCHMEIER, C. 2013. Bace1 and Neuregulin-1 cooperate to control formation and maintenance of muscle spindles. *The European Molecular Biology Organization Journal,* 32**,** 2015-2028.

CITRON, M., OLTERSDORF, T., HAASS, C., MCCONLOGUE, L., HUNG, A. Y., SEUBERT, P., VIGO-PELFREY, C., LIEBERBURG, I. & SELKOE, D. J. 1992. Mutation of the β-amyloid precursor protein in familial Alzheimer's disease increases β-protein production. *Nature,* 360**,** 672-674.

COLE, S. L. & VASSAR, R. 2007. The Basic Biology of BACE1: A Key Therapeutic Target for Alzheimer's Disease. *Current Genomics,* 8**,** 509-530.

CONGREVE, M., AHARONY, D., ALBERT, J., CALLAGHAN, O., CAMPBELL, J., CARR, R. A. E., CHESSARI, G., COWAN, S., EDWARDS, P. D., FREDERICKSON, M., MCMENAMIN, R., MURRAY, C. W., PATEL, S. & WALLIS, N. 2007. Application of Fragment Screening by X-ray Crystallography to the Discovery of Aminopyridines as Inhibitors of B-Secretase. *Journal of Medical Chemistry,* 50**,** 1124-1132.

CORDER, E. H., SAUNDERS, A. M., STRITTMATTER, W. J., SCHMECHEL, D. E., GASKELL, P. C., SMALL, G. W., ROSES, A. D., HAINES, J. L. & PERICAK-VANCE, M. A. 1993. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science,* 261**,** 921-923.

COSTANTINI, C., KO, M. H., JONAS, M. C. & PUGLIELLI, L. 2007. A reversible form of lysine acetylation in the ER and Golgi lumen controls the molecular stabilization of BACE1. *Biochemical Journal,* 407**,** 383-395.

COTMAN, C. W., POON, W. W., RISSMAN, R. A. & BLURTON-JONES, M. 2005. The role of caspase cleavage of tau in Alzheimer disease neuropathology. *J Neuropathol Exp Neurol,* 64**,** 104-12.

CUMMINGS, J., CHO, W., WARD, M., FREISENHAHN, M., BRUNSTEIN, F., HONIGBERG, L., CLAYTON, D., MORTENSEN, D., HO, C. & PAUL, R. 2014. A randomized, double-blind, placebo-controlled phase 2 study to evaluate the efficacy and safety of crenezumab in patients with mild to moderate Alzheimer's disease. *The Journal of the Alzheimer's Association,* 10**,** 275.

DAGERT, M. & EHRLICH, S. D. 1979. Prolonged incubation in calcium chloride improves the competence of Escherichia coli cells. *Gene,* 6**,** 23-28.

DAINA, A., MICHIELIN, O. & ZOETE, V. 2014. iLOGP: a simple, robust, and efficient description of n-octanol/water partition coefficient for drug design using the GB/SA approach. *Journal of Chemical Information and Modeling,* 54**,** 3284-3301.

DE STROOPER, B. 2014. Lessons from a Failed g-Secretase Alzheimer Trial. *Cell,* 4**,** 721-726.

DELANEY, J. S. 2004. ESOL: Estimating aqueous solubility directly from molecular structure. *Journal of Chemical Information and Computer Sciences,* 44**,** 1000-1005.

DISLICH, B. & LICHTENTHALER, S. F. 2012. The membrane-bound aspartyl protease BACE1: molecular and functional properties in Alzheimer's disease and beyond. *Fronteirs in Physiology,* 3**,** 1-16.

DOMINGUEZ, D., TOURNEY, J., HARTMANN, D., HUTH, T., CRYNS, K., DEFORCE, S., SERNEELS, L., CAMACHO, I. E., MARJAUX, E., CRAESSAERTS, K., ROEBROEK, A. J. M., SCHWAKE, M., D'HOOGE, R., BACH, P., KALINKE, U., MOECHARS, D., ALZHEIMER, C., REISS, K., SAFTIG, P. & DE STROOPER, B. 2005. Phenotypical and biochemical analysis of BACE1 and BACE2 deficient mice *The Journal of Biological Chemistry,* 280**,** 30797-30806.

DOODY, R. S., THOMAS, R. G., FARLOW, M., IWATSUBO, T., VELLAS, B., JOFFE, S., KIEBURTZ, K., RAMAN, R., SUN, X., AISEN, P. S., SIEMERS, E., LIU-SIEFERT, H. & MOHS, R. 2014. Phase 3 Trials of Solanezumab for Mild-to-Moderate Alzheimer's Disease. *The New England Journal of Medicine,* 370**,** 311-321.

DOS SANTOS, L. R., PIMASSONI, L. H. S., SENA, G. G. S., CAMPOREZ, D., BELCAVELLO, L., TRANCOZO, M., MORELATO, R. L., ERRERA, F. I. V., BUENO, M. R. P. & DE PAULA, F. 2017. Validating GWAS Variants from Microglial Genes Implicated in Alzheimer's Disease. *Journal of Molecular Neuroscience,* 62**,** 215-221.

DRACHMAN, D. A. 2014. The amyloid hypothesis, time to move on: Amyloid is the downstream result, not cause, of Alzheimer's disease. *Alzheimer's & Dementia,* 10**,** 372-380.

EATON, R. C. & FARLEY, R. D. 1974. Spawning cycle and egg production of zebrafish, *Brachydanio rerio*, in the laboratory. *Copeia,* 1974**,** 195-204.

EATON, R. C. & FARLEY, R. D. 1975. Development of the Mauthner neurons in embryos and larvae of the zebrafish *Brachydanio reio*. *Copeia,* 1973**,** 673-682.

EMMONS, T. L., SHUCK, M. E., BABCOCK, M. S., HOLLOWAY, J. S., LEONE, J. W., DURBIN, J. D., PADDOCK, D. J., PRINCE, D. B., HEINRIKSON, R. L., FISCHER, H. D., BIENKOWSKI, M. J., BENSON, T. E. & TOMASSELLI, A. G. 2008. Large-Scale Purification of Human BACE Expressed in Mammalian Cells and Removal of the Prosegment with HIV-1 Protease to Improve Crystal Diffraction. *Protein and Peptide Letters,* 15**,** 119-130.

ENGELHART, M. J., GEERLINGS, M. I., RUITENBERG, A., VAN SWEITEN, J. C., HOFMAN, A., WITTEMAN, J. C. M. & BRETELER, M. M. B. 2012. Dietary intake of antioxidants and risk of Alzheimer disease. *The Journal of the American Medical Association,* 287**,** 3223-3229.

ESLER, W. P., KIMBERLY, W. T., OSTASZWESKI, B. L., YE, W., DIEHL, T. S., SELKOE, D. J. & WOLFE, M. S. 2002. Activity-dependent isolation of the presenilin–γ-secretase complex reveals nicastrin and a γ substrate. *Proceedings of theNational Academy of Sciences of the United States of America,* 99**,** 2720-2725.

FRIESNER, R. A., BANKS, J. L., MURPHY, R. B., HALGREN, T. A., KLICIC, J. J., MAINZ, D. T., REPASKY, M. P., KNOLL, E. H., SHELLEY, M., PERRY, J. K., SHAW, D. E., FRANCIS, P. & SHENKIN, P. S. 2004. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry,* 47**,** 1739-1749.

GHRIBI, O., LARSEN, B., SCHRAG, M. & HERMAN, M. M. 2006. High cholesterol content in neurons increases BACE, β-amyloid, and phosphorylated tau levels in rabbit hippocampus. *Experimental Neurology,* 200**,** 460-467.

GLENNER, G. G. & WONG, C. W. 1984. Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochemical and biophysical research communications,* 120**,** 885-890.

GOLDRING, J. P. D. & COETZER, T. H. T. 2003. Isolation of chicken immunoglobulins (IgY) from egg yolk. *Biochemistry and Molecular Biology Education,* 31**,** 185-187.

GRIMM, M. O. W., ZINSER, E. G., GRÖSGEN, S., HUNDSDÖRFER, B., ROTHHAAR, T. L., BURG, V. K., KAESTNER, L., BAYER, T. A., LIPP, P., MÜLLER, U., GRIMM, H. S. & HARTMANN, T. 2012. Amyloid Precursor Protein (APP) Mediated Regulation of Ganglioside Homeostasis Linking Alzheimer's Disease Pathology with Ganglioside Metabolism. *PLOS ONE,* 7**,** e34095.

GUO, J. L. & LEE, V. M. Y. 2011. Seeding of Normal Tau by Pathological Tau Conformers Drives Pathogenesis of Alzheimer-like Tangles. *The Journal of Biological Chemistry,* 286**,** 15317-15331.

GUO, J. T., YU, J., GRASS, D., DE BEER, F. C. & KINDY, M. S. 2002. Inflammation-dependent cerebral deposition of serum amyloid a protein in a mouse model of amyloidosis. *J Neurosci,* 22**,** 5900-9.

HAASS, C., LEMERE, C. A., CAPELL, A., CITRON, M., SEUBERT, P., SCHENK, D., LANNFELT, L. & SELKOE, D. J. 1995. The Swedish mutation causes early-ons*et al*zheimer's disease by β-secretase cleavage within the secretory pathway. *Nature Medicine,* 1**,** 1291-1296.

HAWKES, N. 2017. Merck ends trial of potential Alzheimer's drug verubecestat. *The British medical journal,* 356**,** 1.

HAYLEY, M., PERSPICACE, S., SCHULTHESS, T. & SEELIG, J. 2009. Calcium enhances the proteolytic activity of BACE1: An in vitro biophysical and biochemical characterization of the BACE1–calcium interaction. *Biochim. Biophys. Acta - Biomembranes,* 1788**,** 1933-1938.

HE, Y., RUGANZU, J. B., ZHENG, Q., WU, X., JIN, H., PENG, X., DING, B., LIN, C., JI, S., MA, Y. & YANG, W. 2020. Silencing of LRP1 Exacerbates Inflammatory

Response*via*TLR4/NF-κB/MAPKs Signaling Pathways in APP/PS1 Transgenic Mice. *Molecular Neurobiology,* 57**,** 3727-3743.

HEMMING, M. L., ELIAS, J. E., GYGI, S. P. & SELKOE, D. J. 2009. Identification of β-Secretase (BACE1) Substrates Using Quantitative Proteomics. *Proceedings of the National Academy of Sciences of the United States of America,* 4**,** 1-14.

HOPKINS, C. R. 2011. ACS Chemical Neuroscience Molecule Spotlight on ELND006: Another γ-Secretase Inhibitor Fails in the Clinic. *Neuroscience* 2**,** 279-280.

HOWE, K., CLARK, M. D., TORROJA, C. F., TORRANCE, J., BERTHELOT, C., MUFFATO, M., COLLINS, J. E., HUMPHRAY, S., MCLAREN, K., MATTHEWS, L., MCLAREN, S., SEALY, I., CACCAMO, M., CHURCHER, C., SCOTT, C., BARRETT, J. C., KOCH, R., RAUCH, G.-J., WHITE, S., CHOW, W., KILIAN, B., QUINTAIS, L. T., GUERRA-ASSUNÇÃO, J. A., ZHOU, Y., GU, Y., YEN, J., VOGEL, J.-H., EYRE, T., REDMOND, S., BANERJEE, R., CHI, J., FU, B., LANGLEY, E., MAGUIRE, S. F., LAIRD, G. K., LLOYD, D., KENYON, E., DONALDSON, S., SEHRA, H., ALMEIDA-KING, J., LOVELAND, J., TREVANION, S., JONES, M., QUAIL, M., WILLEY, D., HUNT, A., BURTON, J., SIMS, S., MCLAY, K., PLUMB, B., DAVIS, J., CLEE, C., OLIVER, K., CLARK, R., RIDDLE, C., ELLIOTT, D., THREADGOLD, G., HARDEN, G., WARE, D., BEGUM, S., MORTIMORE, B., KERRY, G., HEATH, P., PHILLIMORE, B., TRACEY, A., CORBY, N., DUNN, M., JOHNSON, C., WOOD, J., CLARK, S., PELAN, S., GRIFFITHS, G., SMITH, M., GLITHERO, R., HOWDEN, P., BARKER, N., LLOYD, C., STEVENS, C., HARLEY, J., HOLT, K., PANAGIOTIDIS, G., LOVELL, J., BEASLEY, H., HENDERSON, C., GORDON, D., AUGER, K., WRIGHT, D., COLLINS, J., RAISEN, C., DYER, L., LEUNG, K., ROBERTSON, L., AMBRIDGE, K., LEONGAMORNLERT, D., MCGUIRE, S., GILDERTHORP, R., GRIFFITHS, C., MANTHRAVADI, D., NICHOL, S., BARKER, G., *et al.* 2013. The zebrafish reference genome sequence and its relationship to the human genome. *Nature,* 496**,** 498-503.

HU, X., HICKS, C. W., HE, W., WONG, P., MACKLIN, W. B., TRAPP, B. D. & YAN, R. 2006. Bace1 modulates myelination in the central and peripheral nervous system. *Nature Neuroscience,* 9**,** 1520-1525.

HU, X. T., HE, W., LUO, X., TSUBOTA, K. E. & YAN, R. 2013. BACE1 Regulates Hippocampal Astrogenesis*via*the Jagged1-Notch Pathway. *Cell Reports,* 4**,** 40-49.

HUSE, J. T., PIJAK, D. S., LESLIE, G. J., LEE, V. M. Y. & DOMS, R. W. 2000. Maturation and Endosomal Targeting of BACE: The Alzheimer's Disease ß-Secretase. *The Journal of Biological Chemistry,* 275**,** 33729-33737.

IRWIN, J. J., DUAN, D., TOROSYAN, H., DOAK, A. K., ZIEBART, K. T., STERLING, T., TUMANIAN, G. & SHOICHET, B. K. 2015. An aggregation advisor for ligand discovery. *Journal of Medicinal Chemistry,* 58**,** 7076-7087.

KABIR, A., HONDA, R. P., KAMATARI, Y. O., ENDO, S., FUKUOKA, M. & KUWATA, K. 2016. Effects of ligand binding on the stability of aldo–keto reductases: Implications for stabilizer or destabilizer chaperones. *Protein Science,* 25**,** 2132-2141.

KANG, J., LEMAIRE, H. G., UNTERBACK, A., SALBAUM, J. M., MASTERS, C. L., GRZESCHIK, K. H., MULTHAUP, G., BEYREUTHER, K. & MÜLLER-HILL, B. 1987. The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor. *Nature,* 325**,** 733-736.

KEPP, K. P. 2016. Ten Challenges of the Amyloid Hypothesis of Alzheimer's Disease. *Journal of Alzheimer's Disease,* 55**,** 447-457.

KIM, E.-A., KIM, S.-Y., YE, B.-R., KIM, J., KO, S.-C., LEE, W. W., KIM, K.-N., CHOI, I.-W., JUNG, W.-K. & HEO, S.-J. 2018. Anti-inflammatory effect of Apo-9′-fucoxanthinone*via*inhibition of MAPKs and NF-kB signaling pathway in LPS-stimulated RAW 264.7 macrophages and zebrafish model. *International Immunopharmacology,* 59**,** 339-346.

KIMBERLY, W. T., ESLER, W. P., YE, W., OSTASZWESKI, B. L., GAO, J., DIEHL, T. S., SELKOE, D. J. & WOLFE, M. S. 2003a. Notch and the Amyloid Precursor Protein Are Cleaved by Similar γ-Secretase(s). *Biochemistry,* 42**,** 137-144.

KIMBERLY, W. T., LAVOIE, M. J., OSTASZWESKI, B. L., YE, W., WOLFE, M. S. & SELKOE, D. J. 2003b. γ-Secretase is a membrane protein complex comprised of

presenilin, nicastrin, aph-1, and pen-2. *Proceedings of theNational Academy of Sciences of the United States of America,* 100**,** 6382-6387.

KRAEMER, B. C., ZHANG, B., LEVERENZ, J. B., THOMAS, J. H., TROJANOWSKI, J. Q. & SCHELLENBERG, G. D. 2003. Neurodegeneration and defective neurotransmission in a Caenorhabditis elegans model of tauopathy. *Proceedings of theNational Academy of Sciences of the United States of America,* 100**,** 9980-9985.

KUBINYI, H. 1997. QSAR and 3D QSAR in drug design Part 1: methodology. *Drug Discovery Today,* 2**,** 457-467.

KUHN, P. H., KORONIAK, K., KOGI, S., COLOMBO, A., ZEITSCHEL, U., WILLEM, M., VOLBRACHT, C., SCHEPERS, U., IMHOF, A., HOFFMEISTER, A., HAASS, C., ROBNER, S., BRASE, S. & LICHTENTHALER, S. F. 2012. Secretome protein enrichment identifies physiological BACE1 protease substrates in neurons. *The European Molecular Biology Organization Journal,* 31**,** 3157-3168.

KUHN, P. H., WANG, H., DISLICH, B., COLOMBO, A., ZIETSCHEL, U., ELLWART, J. W., KREMMER, E., ROBNER, S. & LICHTENTHALER, S. F. 2010. ADAM10 is the physiologically relevant, constitutive α-secretase of the amyloid precursor protein in primary neurons. *The European Molecular Biology Organization Journal,* 29**,** 3020-3032.

KUMAR, A., LANA, E., KUMAR, R., LITHNER, C. U. & DARREH-SHORI, T. 2018. Soluble Ab42 Acts as Allosteric Activator of the Core Cholinergic Enzyme Choline Acetyltransferase. *Frontiers in Molecular Neuroscience,* 11**,** 1-11.

KUMAR, A., ROY, S., TRIPATHI, S. & SHARMA, A. 2017. Molecular docking based virtual screening of natural compounds as potential BACE1 inhibitors: 3D QSAR pharmacophore mapping and molecular dynamics analysis. *Journal of Biochemical Structure and Dynamics,* 34**,** 239-249.

KUNKLE, B. W., GRENIER-BOLEY, B., SIMS, R., BIS, J. C., DAMOTTE, V., NAJ, A. C., BOLAND, A., VRONSKAYA, M., VAN DER LEE, S. J., AMLIE-WOLF, A., BELLENGUEZ, C., FRIZATTI, A., CHOURAKI, V., MARTIN, E. R., SLEEGERS, K., BADARINARAYAN, N., JAKOBSDOTTIR, J., HAMILTON-NELSON, K. L., MORENO-GRAU, S., OLASO, R., RAYBOULD, R., CHEN, Y., KUZMA, A. B., HILTUNEN, M., MORGAN, T., AHMAD, S., VARDARAJAN, B. N., EPELBAUM, J., HOFFMANN, P., BOADA, M., BEECHAM, G. W., GARNIER, J.-G., HAROLD, D., FITZPATRICK, A. L., VALLADARES, O., MOUTET, M.-L., GERRISH, A., SMITH, A. V., QU, L., BACQ, D., DENNING, N., JIAN, X., ZHAO, Y., DEL ZOMPO, M., FOX, N. C., CHOI, S.-H., MATEO, I., HUGHES, J. T., ADAMS, H. H., MALAMON, J., SANCHEZ-GARCIA, F., PATEL, Y., BRODY, J. A., DOMBROSKI, B. A., NARANJO, M. C. D., DANIILIDOU, M., EIRIKSDOTTIR, G., MUKHERJEE, S., WALLON, D., UPHILL, J., ASPELUND, T., CANTWELL, L. B., GARZIA, F., GALIMBERTI, D., HOFER, E., BUTKIEWICZ, M., FIN, B., SCARPINI, E., SARNOWSKI, C., BUSH, W. S., MESLAGE, S., KORNHUBER, J., WHITE, C. C., SONG, Y., BARBER, R. C., ENGELBORGHS, S., SORDON, S., VOIJNOVIC, D., ADAMS, P. M., VANDENBERGHE, R., MAYHAUS, M., CUPPLES, L. A., ALBERT, M. S., DE DEYN, P. P., GU, W., HIMALI, J. J., BEEKLY, D., SQUASSINA, A., HARTMANN, A. M., ORELLANA, A., BLACKER, D., RODRIGUEZ-RODRIGUEZ, E., LOVESTONE, S., GARCIA, M. E., DOODY, R. S., MUNOZ-FERNADEZ, C., SUSSAMS, R., LIN, H., FAIRCHILD, T. J., BENITO, Y. A., *et al.* 2019. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nature Genetics,* 51**,** 414-430.

LAEMMLI, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* 227**,** 680-685.

LIM, D., IYER, A., RONCO, V., GROLLA, A. A., CANONICO, P. L., ARONICA, E. & GENAZZANI, A. A. 2013. Amyloid beta deregulates astroglial mGluR5-mediated calcium signaling*via*calcineurin and Nf-kB. *Glia,* 61**,** 1134-1145.

LIPINSKI, C. A., LOMBARDO, F., DOMINY, B. W. & FEENEY, P. J. 2001. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews,* 46**,** 3-26.

LO, M.-C., AULABAUGH, A., JIN, G., COWLING, R., BARD, J., MALAMAS, M. & ELLESTAD, G. 2004. Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Analytical Biochemistry,* 332**,** 153-159.

LONG, D. & YANG, D. 2009. Buffer Interference with Protein Dynamics: A Case Study on Human Liver Fatty Acid Binding Protein. *Biophysical Journal,* 96**,** 1482-1488.

LUNDGREN, J. L., AHMED, S., SCHEDIN-WEISS, S., GOURAS, G. K., WINBLAD, B., TJERNBERG, L. O. & FRYKMAN, S. 2015. ADAM10 and BACE1 are localized to synaptic vesicles. *Journal of Neurochemistry,* 135**,** 606-615.

LUO, Y., BOLON, B., DAMORE, M. A., FITZPATRICK, D., LIU, H., ZHANG, J., YAN, Q., VASSAR, R. & CITRON, M. 2003. BACE1 (β-secretase) knockout mice do not acquire compensatory gene expression changes or develop neural lesions over time. *Neurobiology of Disease,* 14**,** 81-88.

MAES, J., VERLOOY, L., BUENAFE, O. E., DE WITTE, P. A. M., ESQUERRA, C. V. & CRAWFORD, A. D. 2012. Evaluation of 14 organic solvents and carriers for screening applications in zebrafish embryos and larvae. *PLOS ONE,* 7.

MAGSUMOV, T., FATKHUTDINOVA, A., MUKHAMETZYANOV, T. & SEDOV, I. 2019. The Effect of Dimethyl Sulfoxide on the Lysozyme Unfolding Kinetics, Thermodynamics, and Mechanism. *Biomolecules,* 9**,** 547.

MANCINI, F., DE SIMONE, A. & ANDRISANO, V. 2011. Beta-secretase as a target for Alzheimer's disease drug discovery: an overview of in vitro methods for characterization of inhibitors. *Analytical and Bioanalytical Chemistry,* 400**,** 1979-1996.

MARTIN, Y. C. 2005. A bioavailability score. *Journal of Medicinal Chemistry,* 48**,** 3164-3170.

MARTINEZ, N. J., ASAWA, R. R., CYR, M. G., ZAKHAROV, A., URBAN, D. J., ROTH, J. S., WALLGREN, E., KLUMPP-THOMAS, C., COUSSENS, N. P., RAI, G., YANG, S. M., HALL, M. D., MARUGAN, J. J., SIMEONOV, A. & HENDERSON, M. J. 2018. A widely-applicable high-throughput cellular thermal shift assay (CETSA) using split Nano Luciferase. *Scientific Reports,* 8.

MATSUO, K., SHINDO, A., NIWA, A., TABEI, K., AKATSU, H., HASHIZUME, Y., AKIYAMA, H., AYAKI, T., MAKI, T., SAWAMOTO, N., TAKAHASHI, R., OIKAWA, S. & TOMIMOTO, H. 2017. Complement Activation in Capillary Cerebral Amyloid Angiopathy. *Dementia and Geriatric Cognitive Disorders,* 44**,** 343-353.

MATTSON, M. P., ENGLE, M. G. & RYCHLIK, B. 1991. Effects of Elevated Intracellular Calcium Levels on the Cytoskeleton and Tau in Cultured Human Cortical Neurons. *Molecular and Chemical Neuropathology,* 15**,** 117-142.

MENZEN, T. & FRIESS, W. 2012. High-throughput melting-temperature analysis of a monoclonal antibody by differential scanning fluorimetry in the presence of surfactants. *Journal of Pharmaceutical Sciences,* 102**,** 415-428.

MERCHED, A., XIA, Y., VISVIKIS, S., SEROT, J. M. & SIEST, G. 2000. Decreased high-density lipoprotein cholesterol and serum apolipoprotein AI concentrations are highly correlated with the severity of Alzheimer's disease☆. *Neurobiology of Aging,* 21**,** 27-30.

MIN, Z., TANG, Y., HU, X. T., ZHU, B. L., MA, Y. L., ZHA, J. S., DENG, X. J., YAN, Z. & CHEN, C. J. 2018. Cosmosiin Increases ADAM10 Expression*via*Mechanisms Involving 5'UTR and PI3K Signaling. *Frontiers in Molecular Neuroscience,* 11**,** 1-12.

MODARRESI, F., FAGHIHI, M. A., PATEL, N. S., SAHAGAN, B. G., WHALESTEDT, C. & LOPEZ-TOLEDANO, M. A. 2011. Knockdown of BACE1-AS Nonprotein-Coding Transcript Modulates Beta-Amyloid-Related Hippocampal Neurogenesis. *International Journal of Alzheimer's Disease,* 2011**,** 1-11.

MOLINA, D. M., JAFARI, R., IGNATUSHCHENKO, M., SEKI, T., LARSSON, E. A., CHEN, D., SREEKUMAR, L., CAO, Y. & NORDLUND, P. 2013. Monitoring Drug Target Engagement in Cells and Tissues Using the Cellular Thermal Shift Assay. *Science,* 341**,** 84-87.

MORI, H., KONDO, J. & IHARA, Y. 1987. Ubiquitin is a component of paired helical filaments in Alzheimer's disease. *Science,* 235**,** 1641-1644.

MURRAY, C. W. & REES, D. C. 2009. The rise of fragment-based drug discovery. *Nature, Chemistry,* 1**,** 187-192.

NENISKYTE, U., FRICKER, M. & BROWN, G. C. 2016. Amyloid β induces microglia to phagocytose neurons*via*activation of protein kinase Cs and NADPH oxidase. *The International Journal of Biochemistry & Cell Biology,* 81**,** 346-355.

NIESEN, F. H., BERGLUND, H. & VEDADI, M. 2007. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols,* 2**,** 2212-2221.

NUKINA, N. & IHARA, Y. 1986. One of the antigenic determinants of paired helical filaments is related to tau protein. *Journal of biochemistry,* 99**,** 1541-1544.

ÖBERG, F., SJÖHAMN, J., FISCHER, G., MOBERG, A., PEDERSEN, A., NEUTZE, R. & HEDFALK, K. 2011. Glycosylation increases the thermostability of human aquaporin 10 protein. *J. Biol. Chem,* 286**,** 31915-31923.

OECD 2013. OECD Guidelines for the testing of chemicals. Section 2: effects on biotic systems test no. 236: Fish embryo acute toxicity (FET) test.

OTT, A., SLOOTER, A. J. C., HOFMAN, A., VAN HARSKAMP, F., WITTEMAN, J. C. M., VAN BROECKHOVEN, C., VAN DUJIN, C. M. & BRETELER, M. M. B. 1998. Smoking and risk of dementia and Alzheimer's disease in a population-based cohort study: the Rotterdam Study. *The Lancet,* 351.

PANZICA-KELLY, J. M., ZHANG, C. X., DANBERRY, T. L., FLOOD, A., DELAN, J. W., BRANNEN, K. C. & AUGUSTINE-RAUCH, K. A. 2010. Morphological score assignment guidelines for the dechorionated zebrafish teratogenicity assay. *Developmental and Reproductive Toxicology,* 89**,** 382-395.

PAUL, T. J., BARMAN, A., OZBIL, M., BORA, R. P., ZHANG, T., SHARMA, G., HOFFMANN, Z. & PRABHAKAR, R. 2016. Mechanisms of peptide hydrolysis by aspartyl and metalloproteases. *Physical Chemistry Chemical Physics,* 18**,** 24790-24801

PIGONI, M., WANNGREN, J., KUHN, P. H., MUNRO, K. M., GUNNERSEN, J. M., TAKESHIMA, H., FEEDERLE, R., VOYTYUK, I., DE STROOPER, B., LEVASSEUR, M. D., HRUPKA, B. J., MULLER, S. A. & LICHTENTHALER, S. F. 2016. Seizure protein 6 and its homolog seizure 6-like protein are physiological substrates of BACE1 in neurons. *Molecular Neurodegeneration,* 11**,** 1-18.

PIRTTIMAKI, T. M., CODADU, N. K., AWNI, A., PRATIK, P., NAGEL, D. A., HILL, E. J., DINELEY, K. T. & PARRI, H. R. 2013. α7 Nicotinic Receptor-Mediated Astrocytic Gliotransmitter Release: Aβ Effects in a Preclinical Alzheimer's Mouse Model. *Public Library of Science One,* 8**,** 1-12.

POSTINA, R., SCHROEDER, A., DEWACHTER, I., BOHL, J., SCHMITT, U., KOJRO, E., PRINZEN, C., ENDRES, K., HIEMKE, C., BLESSING, M., FLAMEZ, P., DEQUENNE, A., GODAUX, E., VAN LEUVEN, F. & FAHRENHOLZ, F. 2004. A disintegrin-metalloproteinase prevents amyloid plaque formation and hippocampal defects in an Alzheimer disease mouse model. *Journal of Clinical Investigation,* 113**,** 1456-1464.

PRETE, D. D., CHECLER, F. & CHAMI, M. 2014. Ryanodine receptors: physiological function and deregulation in Alzheimer disease. *Molecular Neurodegeneration,* 9.

RASOOL, C. G., SVENDSEN, C. N. & SELKOE, D. J. 1986. Neurofibrillary degeneration of cholinergic and noncholinergic neurons of the basal forebrain in Alzheimer's disease. *Annals of Neurology,* 20**,** 482-488.

REYNOLDS, I. J. & HASTINGS, T. G. 1995. Glutamate induces the production of reactive oxygen species in cultured forebrain neurons following NMDA receptor activation. *Journal of Neuroscience,* 15**,** 3318-3327.

ROBINSON, M. J., MATEJTCHUK, P., BRISTOW, A. F. & DALBY, P. A. 2018. Tm-values and unfolded fraction can predict aggregation rates for granulocyte colony stimulating factor variant formulations but not under predominantly native conditions. *Molecular pharmaceutics,* 15**,** 256-267.

ROBNER, S., LANGE-DOHNA, C., ZEITSCHEL, U. & PEREZ-POLO, J. R. 2004. Alzheimer's disease β-secretase BACE1 is not a neuron-specific enzyme. *Journal of Neurochemistry,* 92**,** 226-234.

ROSES, A. D. 2006. Pharmacogenetics in drug discovery and development: a translational perspective. *Nature Reviews, Drug Discovery,* 7**,** 807-817.

RUSHWORTH, J. V., GRIFFITHS, H. H., WATT, N. T. & HOOPER, N. M. 2013. Prion Protein-mediated Toxicity of Amyloid-β Oligomers Requires Lipid Rafts and the Transmembrane LRP1. *Journal of Biological Chemistry,* 288**,** 8935-8951.

SARDANA, V., XU, B., ZUGAY-MURPHY, J., CHEN, Z., SARDANA, M., DARKE, P. L., MUNSHI, S. & KUO, L. C. 2003. A general procedure for the purification of human b-secretase expressed in Escherichia coli. *Protein Expression And Purification,* 34**,** 190-196.

SASAKI, A., SHOJI, M., HARIGAYA, Y., KAWARABAYASHI, T., IKEDA, M., NAITO, M., MATSUBARA, E., ABE, K. & NAKAZATO, Y. 2002. Amyloid cored plaques in Tg2576 transgenic mice are characterized by giant plaques, slightly activated microglia, and the lack of paired helical filament-typed, dystrophic neurites. *Virchows Archiv: European Journal of Pathology,* 441**,** 358-367.

SATHYA, M., PREMKUMAR, P., KARTHICK, C., MOORTHI, P., JAYACHANDRAN, K. S. & ANUSUYADEVI, M. 2012. BACE1 in Alzheimer's disease. *Clinica ChimicaActa,* 414**,** 171-178.

SCHAEFER, P. M., VON EINEM, B. V., WALTHER, P., CALZIA, E. & CVON ARNIM, C. A. F. 2016. Metabolic characterization of intact cells reveals intracellular amyloid beta but not its precursor protein to reduce mitochondrial respiration. *PLOS ONE,* 11.

SCHMECHEL, A., STRAUSS, M., SCHLICKSUPP, A., PIPKORN, R., HAASS, C., BAYER, T. A. & MULTHAUP, G. 2004. Human BACE Forms Dimers and Colocalizes with APP. *The Journal of Biological Chemistry,* 279**,** 39710-39717.

SCHNEIDER, C. A., RASBAND, W. S. & ELICEIRI, K. W. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods,* 9**,** 671-675.

SCOTT, J. D., LI, S. W., BRUNSKILL, A. P. J., CHEN, X., COX, K., CUMMING, J. N., FORMAN, M., GILBERT, E. J., HODGSON, R. A., HYDE, L. A., JIANG, Q., ISERLOH, U., KAZAKEVICH, I., KUVELKAR, R., MEI, H., MEREDITH, J., MISIASZEK, J., ORTH, P., ROSSITER, L. M., SLATER, M., STONE, J., STRICKLAND, C. O., VOIGT, J. H., WANG, G., WANG, H., WU, Y., GREENLEE, W. J., PARKER, E. M., KENNEDY, M. E. & STAMFORD, A. W. 2016. Discovery of the 3-Imino-1,2,4-thiadiazinane 1,1-Dioxide derivative verubecestat (MK-8931)–a β-site amyloid precursor protein cleaving enzyme 1inhibitor for the treatment of Alzheimer's disease. *J. Med. Chem,* 59**,** 10435-10450.

SEVIGNY, J., CHIAO, P., BUSSIERE, T., WEINREB, P. H., WILLIAMS, L., MAIER, M., DUNSTAN, R., SALLOWAY, S., CHEN, T., LING, Y., O'GORMAN, J., QIAN, F., ARASTU, M., LI, M., CHOLLATE, S., BRENNAN, M. S., QUINTERO-MONZON, O., SCANNEVIN, R. H., ARNOLD, H. M., ENGBER, T., RHODES, K., FERRERO, J., HANG, Y., MIKULSKIS, A., GRIMM, J., HOCK, C., NITSCH, R. M. & SANDROCK, A. 2016. The antibody aducanumab reduces Aβ plaques in Alzheimer's disease. *Nature,* 50**,** 50-56.

SHANKAR, G. M., BLOODGOOD, B. L., TOWNSEND, M., WALSH, D. M., SELKOE, D. J. & SABATINI, B. L. 2007. Natural Oligomers of the Alzheimer Amyloid-β Protein Induce Reversible Synapse Loss by Modulating an NMDA-Type Glutamate Receptor-Dependent Signaling Pathway. *Neurobiology of Disease,* 27**,** 2866-2875.

SHERRINGTON, R., FROELICH, S., SORBI, S., CAMPION, D., CHI, H., ROGAEVA, E. A., LEVESQUE, G., ROGAEV, E. I., LIN, C., LIANG, Y., IKEDA, M., MAR, L., BRICE, A., AGID, Y., PERCY, M. E., CLERGET-DARPOUX, F., PIACENTINI, S., MARCON, G., NACMIAS, B., AMADUCCI, L., FREBOURG, T., LANNFELT, L., ROMMENS, J. M. & ST GEORGE-HYSLOP, P. H. 1996. Alzheimer's Disease Associated with Mutations in Presenilin 2 is Rare and Variably Penetrant. *Human molecular genetics,* 5**,** 985-988.

SKOVRONSKY, D. M., MOORE, D. B., MILLA, M. E., DOMS, R. W. & LEE, V. M. Y. 2000. Protein Kinase C-dependent α-Secretase Competes with β-Secretase for Cleavage of Amyloid-β Precursor Protein in the Trans-Golgi Network. *The Journal of Biological Chemistry,* 275**,** 2568-2575.

SORRELL, F. J., GREENWOOD, G. K., BIRCHALL, K. & CHEN, B. 2010. Development of a differential scanning fluorimetry based high throughput screening assay for the

discovery of affinity binders against an anthrax protein. *Journal of Pharmaceutical and Biomedical Analysis,* 52**,** 802-808.

STOCKLEY, J. H. & O'NEILL, C. 2007. The proteins BACE1 and BACE2 and β-secretase activity in normal and Alzheimer's disease brain. *Biochemical Society Transactions,* 35**,** 574-576.

SU, B., WANG, X., LEE, H. G., TABATON, M., PERRY, G., SMITH, M. A. & ZHU, X. 2010. Chronic oxidative stress causes increased tau phosphorylation in M17 neuroblastoma cells. *Neuroscience Letters,* 468**,** 267-271.

TABET, N. 2006. Acetylcholinesterase inhibitors for Alzheimer's disease: anti-inflammatories in acetylcholine clothing! *Age and Ageing,* 35**,** 336-338.

TAM, E. M., MOORE, T. R., BUTLER, G. S. & OVERALL, C. M. 2004. Characterization of the Distinct Collagen Binding, Helicase and Cleavage Mechanisms of Matrix Metalloproteinase 2 and 14 (Gelatinase A and MT1-MMP). *Journal of Biological Chemistry,* 279**,** 43336-43344.

TEAME, T., ZHANG, Z., RAN, C., ZHANG, H., YANG, Y., DING, Q., XIE, M., GAO, C., YE, Y., DUAN, M. & ZHOU, Z. 2019. The use of zebrafish (Danio rerio) as biomedical models. *Animal Frontiers,* 9**,** 68-77.

THOMAS, R. J. 1975. The toxicologic and teratologic effects of Δ9-tetrahydrocannabinol in the Zebrafish embryo. *Toxicology and Applied Pharmacology,* 32**,** 184-190.

TIWARI, M. K. & KEPP, K. P. 2015. B-Amyloid pathogenesis: Chemical properties versus cellular levels. *Alzheimer's & Dementia,* 12**,** 184-194.

TOMASSELLI, A. G., PADDOCK, D. J., EMMONS, T. L., MILDNER, A. M., LEONE, J. W., LULL, J. M., CIALDELLA, J. I., PRINCE, D. B., FISCHER, H. D., HEINRIKSON, R. L. & BENSON, T. E. 2008. High yield expression of human BACE constructs in Eschericia coli for refolding, purification, and high resolution diffracting crystal forms. *Protein and Peptide Letters,* 15**,** 131-143.

TOMASSELLI, A. G., QAHWASH, I., EMMONS, T. L., LU, Y., LEONE, J. W., LULL, J. M., FOK, K. F., BANNOW, C. A., SMITH, C. W., BIENKOWSKI, M. J., HEINRIKSON, R. L. & YAN, R. 2003. Employing a superior BACE1 cleavage sequence to probe cellular APP processing. *Journal of Neurochemistry,* 84**,** 1006-1017.

TURNER, R. T., KOELSCH, G., HONG, L., CASTANHEIRA, P., GHOSH, A. & TANG, J. 2001. Subsite Specificity of Memapsin 2 (B-Secretase): Implications for Inhibitor Design. *Biochemistry,* 40**,** 10001-10006.

VASSAR, R., BENNETT, B. D., BABU-KHAN, S., KAHN, S., MENDIAZ, E. A., DENIS, P., TEPLOW, D. B., ROSS, S., AMARANTE, P., LOELOFF, R., LUO, Y., FISHER, S., FULLER, J., EDENSON, S., LILE, J., JAROSINSKI, M. A., BIERE, A. L., CURRAN, E., BURGESS, T., LOUIS, J. C., COLLINS, F., TREANOR, J., ROGERS, G. & CITRON, M. 1999. B-Secretase Cleavage of Alzheimer's Amyloid Precursor Protein by the Transmembrane Aspartic Protease BACE. *Science,* 286**,** 735-741.

VETRIVEL, K. S., BARMAN, A., CHEN, Y., NGUYEN, P. D., WAGNER, S. L., PRABHAKAR, R. & THINAKARAN, G. 2011. Loss of cleavage at β'-site contributes to apparent increase in β-amyloid peptide (Aβ) secretion by β-secretase (BACE1)-glycosylphosphatidylinositol (GPI) processing of amyloid precursor protein. *Journal of biochemistry,* 286**,** 26166-26177.

VETRIVEL, K. S., MECKLER, X., CHEN, Y., NGUYEN, P. D., SEIDAH, N. G., VASSAR, R., WONG, P. C., FUKATA, M., KOUNNAS, M. Z. & THINAKARAN, G. 2009a. Alzheimer disease Abeta production in the absence of S-palmitoylation-dependent targeting of BACE1 to lipid rafts. *Journal of biochemistry,* 284**,** 3793-3803.

VETRIVEL, K. S., MECKLER, X., CHEN, Y., NGUYEN, P. D., SEIDAH, N. G., VASSAR, R., WONG, P. C., FUKATA, M., KOUNNAS, M. Z. & THINAKARAN, G. 2009b. Alzheimer Disease Aβ Production in the Absence of S-Palmitoylation-dependent Targeting of BACE1 to Lipid Rafts. *Journal of Biological Chemistry,* 284**,** 3793-3803.

WANG, W. Y., TAN, M. S., YU, J. T. & TAN, L. 2015. Role of pro-inflammatory cytokines released from microglia in Alzheimer's disease. *Annals of Translational Medicine,* 3**,** 1-15.

WESTMEYER, G. G., WILLEM, M., LICHTENTHALER, S. F., LURMAN, G., MULTHAUP, G., ASSFALG-MACHLEIDT, I., REISS, K., SAFTIG, P. & HAASS,

C. 2004. Dimerization of β-site β-amyloid precursor protein-cleaving enzyme. *J. Biol. Chem,* 279**,** 53205-53212.

WILDMAN, S. A. & CRIPPEN, G. 1999. Prediction of physicochemical parameters by atomic contributions. *Journal of Chemical Information and Computer Sciences,* 39**,** 868-873.

WILKINSON, B. L. & LANDRETH, G. E. 2006. The microglial NADPH oxidase complex as a source of oxidative stress in Alzheimer's disease. *Journal of Neuroinflammation,* 3**,** 1-12.

WRIGHT, T. A., STEWART, J. M., PAGE, R. C. & KONKOLEWICZ, D. 2017. Extraction of Thermodynamic Parameters of ProteinUnfolding Using Parallelized Differential Scanning Fluorimetry. *The Journal of Physical Chemistry Letters,* 8**,** 553-558.

XU, Y., LI, M., GREENBLATT, H., CHEN, W., PAZ, A., DYM, O., PELEG, Y., CHEN, T., SHEN, X., HE, J., JIANG, H., SILMAN, I. & SUSSMAN, J. L. 2012. Flexibility of the flap in the active site of BACE1 as revealed by crystal structures and molecular dynamics simulations. *Acta Crystallographica,* 68**,** 13-25.

YAN, R. 2017. Physiological Functions of the β-Site Amyloid Precursor Protein Cleaving Enzyme 1 and 2. *Frontiers in Molecular Neuroscience,* 10**,** 1-12.

YAN, R., HAN, P., MIAO, H., GREENGARD, P. & XU, H. 2001a. The transmembrane domain of the Alzheimer's β-secretase (BACE1) determines its late Golgi localization and access to APP substrate. *The Journal of Biological Chemistry,* 276**,** 36788-36796.

YAN, R., MUNZNER, J. B., SHUCK, M. E. & BIENKOWSKI, M. J. 2001b. BACE2 Functions as an Alternative α-Secretase in Cells. 276**,** 34019-34027.

YANG, Z., KUBOYAMA, T. & TOHDA, C. 2019. Naringenin promotes microglial M2 polarization and Aβ degradation enzyme expression. *Phytotherapy Research,* 33**,** 1114-1121.

YU, Y., WANG, J., SHAO, Q., SHI, J. & ZHU, W. 2016. The effects of organic solvents on the folding pathway and associated thermodynamics of proteins: a microscopic view. *Scientific Reports,* 6.

ZHANG, L., SONG, L., TERRACINA, G., LIU, Y., PRAMANIK, B. & PARKER, E. 2001. Biochemical Characterization of the γ-Secretase Activity That Produces β-Amyloid Peptides. *Biochemistry,* 40**,** 5049-5055.

APPENDIX A

03 June 2020

Prof Carola Ulrike Niesler (642424)
School of Life Sciences
Pietermaritzburg

Dear Prof Nielser,

Protocol reference number: AREC/029/019
Project title: Establishment of Zebrafish models for the study of human disease and toxicology.

Full Approval – Research Application

With regards to your revised application received on 12 May 2020, the documents submitted have been accepted by the Animal Research Ethics Committee and **FULL APPROVAL** for the protocol has been granted.

Please note: Any Veterinary and Para-Veterinary procedures must be conducted by a SAVC registered VET or SAVC authorized person.

Any alteration/s to the approved research protocol, i.e Title of Project, Location of the Study, Research Approach and Methods must be reviewed and approved through the amendment/modification prior to its implementation. In case you have further queries, please quote the above reference number.

Please note: Research data should be securely stored in the discipline/department for a period of 5 years.

The ethical clearance certificate is only valid for a period of one year from the date of issue. Renewal for the study must be applied for before 02 June 2021.

Attached to the Approval letter is a template of the Progress Report that is required at the end of the study, or when applying for Renewal (whichever comes first). An Adverse Event Reporting form has also been attached in the event of any unanticipated event involving the animals' health / wellbeing.

I take this opportunity of wishing you everything of the best with your study.

Yours faithfully

....
Dr Sanil D Singh, PhD
Chair: Animal Research Ethics Committee

/kr

Figure A.1: Ethics approval (AREC/029/019) was granted for the project entitled: "Establishment of Zebrafish models for the study of human disease and toxicology" on 3 June 2020.12.03
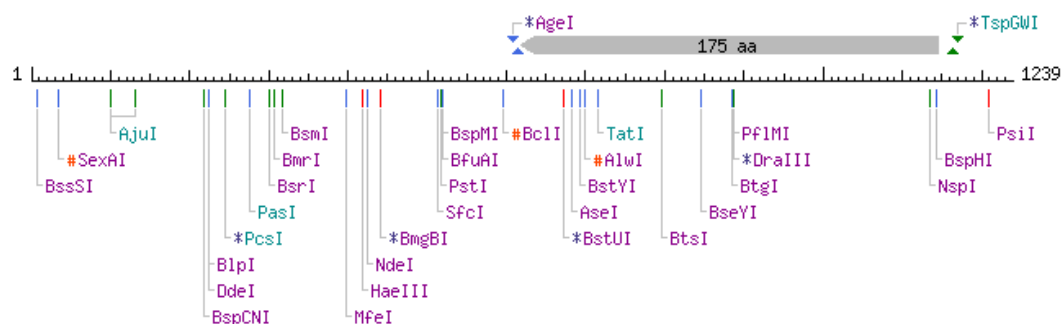
Figure B.1: Map of restriction enzyme cleavage sites found in the BACE1 insert derived from GenBank (GeneID: 23621) using the online NEBCutter tool (New England Biolabs, USA).

APPENDIX C

Table C.1: Analysis of protease cleavage sites found in the BACE1 peptide sequence carried out using the online PeptideCutter tool (Swiss Institute of Bioinformatics, Switzerland).

Protease	Number of Cleavage sites
Arg-C proteinase	25
Asp-N endopeptidase	25
Asp-N endopeptidase + N-terminal Glu	52
BNPS-Skatole	10
Caspase1	1
Caspase2	0
Caspase3	0
Caspase4	0
Caspase5	0
Caspase6	0
Caspase7	0
Caspase8	0
Caspase9	0
Caspase10	0
Chymotrypsin-high specificity (C-term to [FYW], not before P)	45
Chymotrypsin-low specificity (C-term to [FYWML], not before P)	107
Clostripain	25
CNBr	14
Enterokinase	0
GranzymeB	0
Factor Xa	0

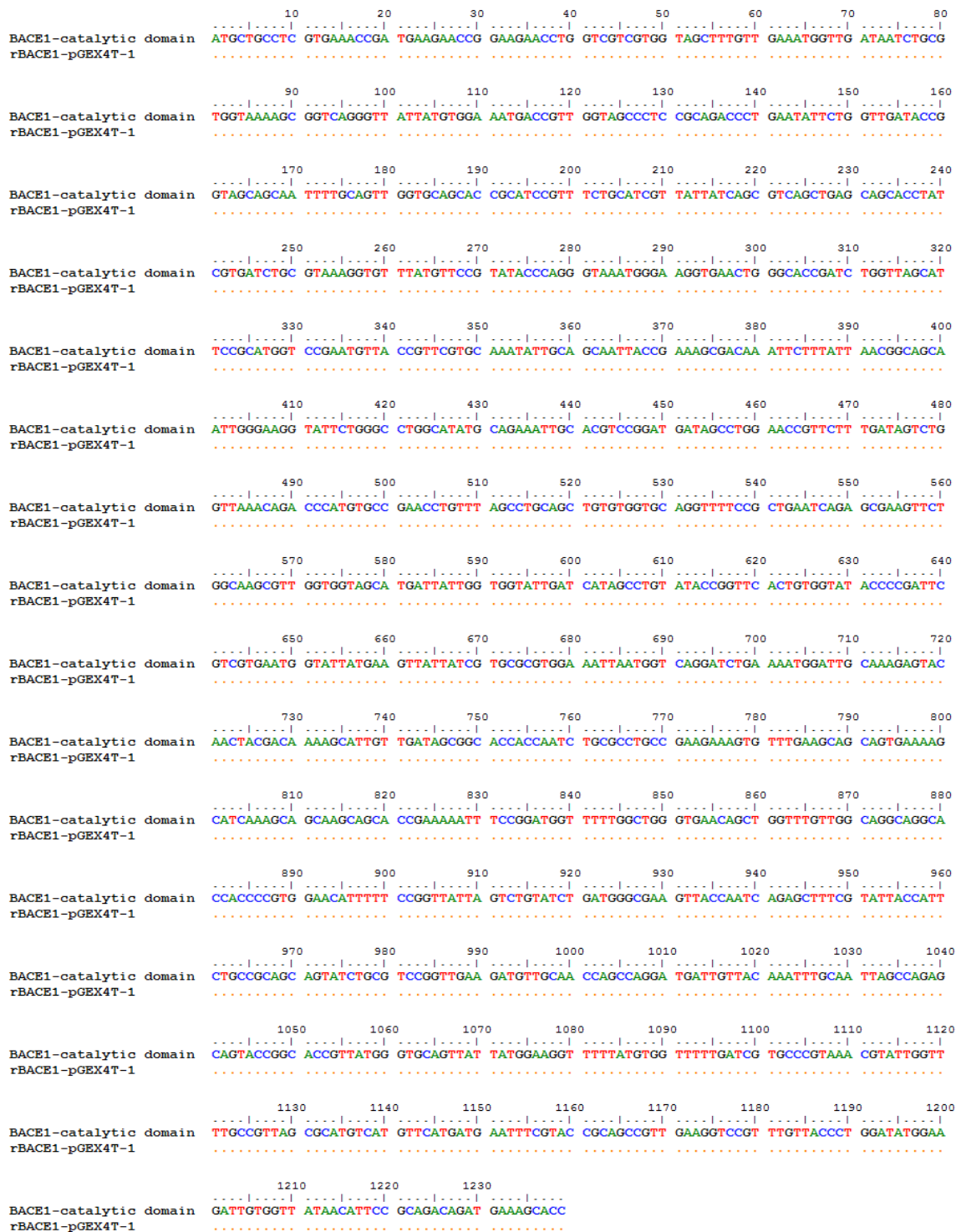|	
Formic acid	25
Glutamyl endopeptidase	27
Hydroxylamine	2
Iodosobenzoic acid	10
LysC	11
LysN	11
NTCB (2-nitro-5-thiocyanobenzoic acid)	11
Pepsin (pH1.3)	96
Pepsin (pH>2)	127
Proline-endopeptidase	3
Proteinase K	239
Staphylococcal peptidase I	25
Tobacco etch virus protease	1
Thermolysin	140
Thrombin	0
Trypsin	34

Figure D.1: Sequence alignment of sequenced BACE1 insert against human BACE1 gene with identical nucleotides marked as dots.
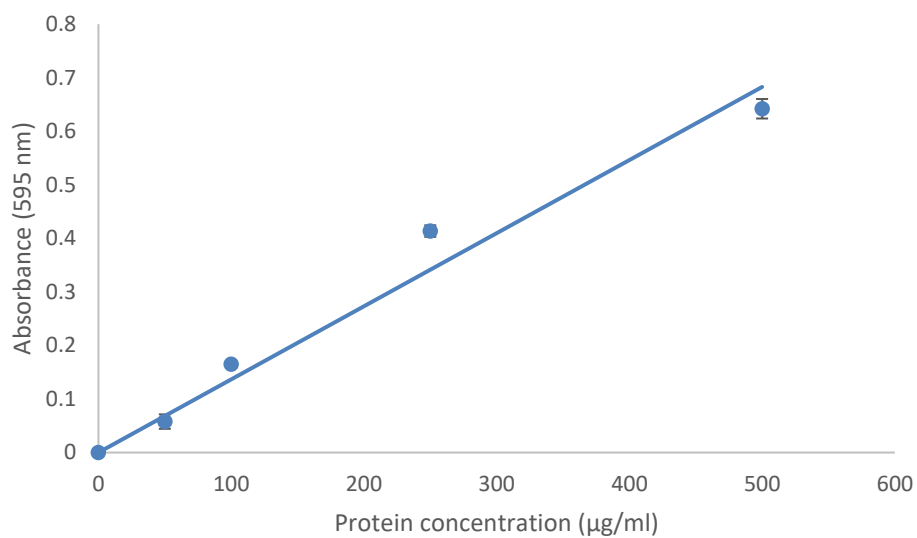
Figure E.1: Reference standard curve for the Bradford assay. Bovine serum albumin standards ranging from 50 to 500 µg/ml were mixed 1:9 with Bradford reagent. Absorbance at 595 nm was measured. The equation for the line of best fit is y = 0.0014x. Error bars represent standard deviations of triplicate samples.
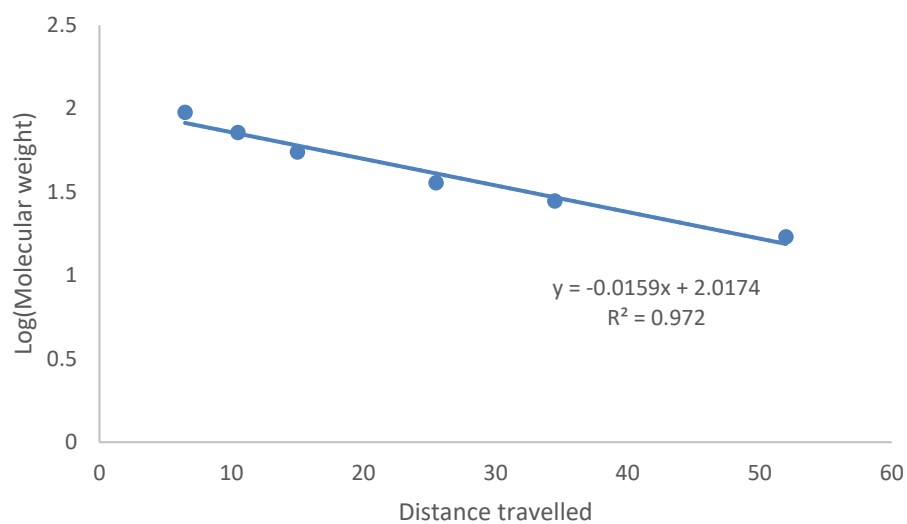
Figure F.1: Reference standard curve for the calculation of molecular weights of proteins separated using sodium dodecyl sulphate gel electrophoresis. Generated by measuring the distance travelled by each protein in the molecular weight marker, plotted against the log(molecular weight) of each protein band.